THINK INDIA JOURNAL

ISSN:0971-1260
Vol-22-Issue-10-November-2019

Analysing Binge-Watching Behaviour Using Data Mining Techniques

[1]Dr K B Priya Iyer, [2]Nishkala G, [2]Priyankha V, [2]Srilalitha R
[1] Associate Professor, Dept. of Computer Science
[2] Under Graduate Student, BCA, Dept. of Computer Science
M.O.P. Vaishnav College for Women (Autonomous), Chennai, India

ABSTRACT

*Advancements in new technology have lead people towards Netflix and Prime wherein people can watch several series and movies sitting at their own convenient time and place These have further lead to a new term "**binge watching**" where people tend to sit for hours to watch a series or movie in rapid succession. Most people consider binge watching as a hobby while for some it becomes a daily routine. The term binge refers to excess eating or drinking. People tend to feel happy when they are binge watching as it gives them a pleasure and relaxation to mind amidst their busy schedules. Binge watching has its own pros and cons which will be discussed in this paper. This paper discusses the percentage of people addicted to binge watching and the effects and causes of binge watching using data mining techniques. Experimental results showed the impact of binge watching is higher among the people of age group 15-27.*

***Keywords**: Binge-watching, data mining, random forest, naïve bayes.*

I. INTRODUCTION

Binge-watching is a habit of viewing television for an extensive time width, generally a sole television show. Binge-watching noticed cultural sensation has become familiar with the support of visual treats streaming programs on channels like Netflix, Amazon video, Hulu, etc. by way of which a person can watch series, shows, and movies on request. Contrarily, more recurring appointment viewing of TV viewing was operated by viewing potency and matured age[12].

This study is about how binge-viewing is affecting people at different ages, what and all the activities people do to binge-watch, how many hours do they spend on binge-watching, etc. by conducting surveys and analyzing the flow of graph through various aspects and classification. Basically, this analysis focuses on the people between the age group from 15-27, to get the outcome of how the young bloods are getting distracted through marathon-viewing. More marathon viewing happened among people not intricate in athletics to move time but not to gather information[13]. Binge watching can be an incredible method to unwind and de-stress, yet it can without much of a stretch become an issue when you consistently organize it over other significant exercises [2].

Like betting and other conduct addictions, binge watching initiates the piece of the mind in charge of "remunerate" capacities, delivering dopamine and making us feel better. After some time, however, our cerebrums produce less dopamine from a similar degree of movement as we develop a degree of resistance. It takes increasingly more of a similar action to give us that equivalent sentiment of happiness, making binge watching that a lot harder to stop [3].

ISSN:0971-1260
Vol-22-Issue-10-November-2019

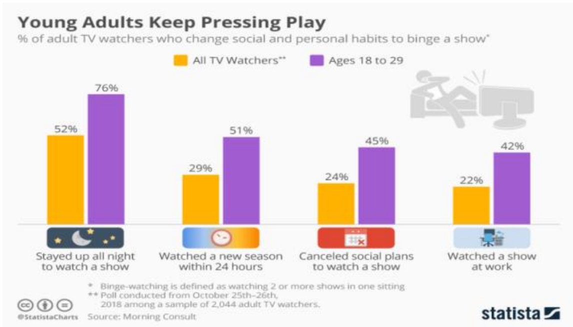


Figure 1: Percentage of Binge watchers who change social and personal habits.

From fig 1, the chart shows the percentage of Binge-watchers which illustrates the people around 18-29 age are more involved in watching shows in different aspects. in the society, who identified themselves as a marathon viewer is 80.6%, and 20.2% had marathon viewed minimum of few times of week[9].

Binge watching shows may really be less critical and less pleasant than observing new scenes on a week by week premise [4]. Specialists found that binge watching shows diminished members' capacity to recollect subtleties from the shows, and binge watchers likewise lower happiness levels from watching the show than the individuals who viewed a similar arrangement just once every day.

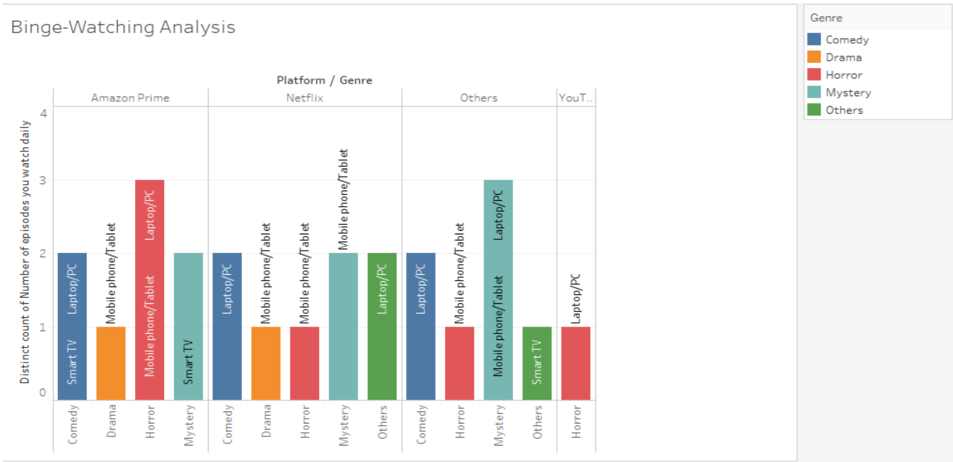
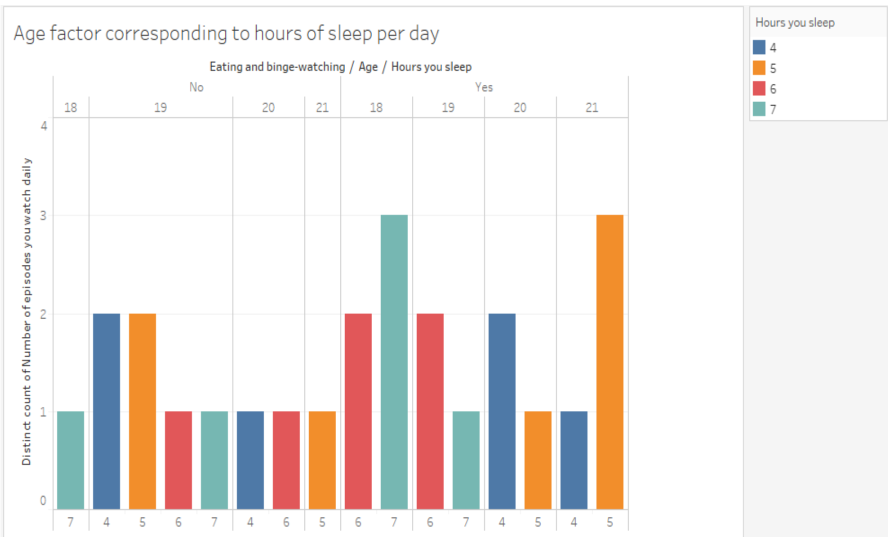


Figure 2: Binge watching corresponding to genre and the device that is mostly used.

As per the survey collected, It obviously portrays 70 percent of its streams end up on Advanced mobile phones and PCs rather than television, tablets[14].



ISSN:0971-1260
Vol-22-Issue-10-November-2019

Figure 3: Age factor corresponding to the hours of sleep per day.

While those Binge watching platforms are making it simpler to purchase a membership from your sofa, in any case it will be simpler to do it from a mobile phone, which records for 40 percent of signups. Laptop/PCs represent another 30 percent. Study research has concluded that TV watching is corresponding with low mental and physical results[10].

From the second figure, we can infer that in case you're staying there binge watching for six hours, in addition to the fact that you are stationary, you're most likely eating whatever's before you and not considering it. Teenagers around age 19 are most likely to binge watch during nights depriving sleep. On an average, they tend to sleep for 3 hours a day which is detrimental to physical and mental health. Above that, the investigation supplements research demonstrating a solid connection between binge viewing and undesirable eating practices. Studies have discovered individuals in general consume unhealthy food while binge watching, however they devour a greater amount of it while binge watching.

II. LITERATURE REVIEW

On February 2014, a survey conducted by Netflix, which comes out with conclusion of 73% of mankind explain binge-watching as “viewing approximately 2-6 episodes or more of a particular TV show in one sitting. There are various names for the term binge-watching, like binge-viewing alternatively marathon-viewing[6].

Recently, a research which was conducted by University of Texas at Austin obtained the result saying, binge-watching television is harmonized with pessimism, isolation, self-moving deficiency and plumpness. The consequence of binge-watching motivation encourage true binge viewing behaviour [16]. Further, the author concluded by saying that even though people compromise, marathon-viewing is harmless addiction, it is not the way good to be watched for a prolonged period of time [1].

Multiple analytical reports have mentioned more than 60 percent of leading channel subscribers binge-view on daily basis. It was already found from linked binge-watching to a bad sleep practice, escalation of insomnia and fatigue and also proven facts stating marathon-viewing could edge to a developing cognitive alertness, thence impact on sleep[7]. According to a study research, binge watching better conveys an escape fulfillment for people than meeting viewing.[8].

III. DATASET DESCRIPTION

The data is collected among different age groups. Number of respondents are 201. The data was collected based on their age, weekly study time, category, platform, frequency of watching, sleep hours, extra activities and number of episodes.

ATTRIBUTE	DESCRIPTION
AGE	Millennial (12 - 25 years)
WST	Weekly study time (in hrs)
CATEGORY	Movies, Series.etc.,
PLATFORM	Netflix, Hulu, Hotstar.etc.,
FREQUENCY	Once a week, Regular.etc.,
SLEEP_HOURS	Duration of sleep (in hrs)
EXTRA_ACTS	Activities (Yes or No)
NO_OF_EPI	Number of episodes (2 episodes, 3-5 episodes, Entire season)

Table 1: Dataset Attributes and their description

IV DATA ANALYSIS

NAÏVE BAYES CLASSIFICATION

ISSN:0971-1260
Vol-22-Issue-10-November-2019

Naïve Bayes-Model:

This model is very much useful for predictive analysis. Naive Bayes classifier is a classification ability based on Bayes' Theorem with a presupposition of independence amid predictors. It presumes that the existence of a specific feature in a collection of objects, unrelated to the existence of any other feature.

This model assists us to obtain the inference that youngsters at the average age group from 15 to 20 binge view regularly with sacrificing their sleep nearly up to 3-5 hours. Their duration of sleep is approximately 6 hours per day. Few of the people binge view regularly, some participate more in extracurricular activities, and few does both regularly. . Naïve Bayes classifier with homogenous polynomial, Bernoulli and Gaussian incident models are not completely Bayesian [18].

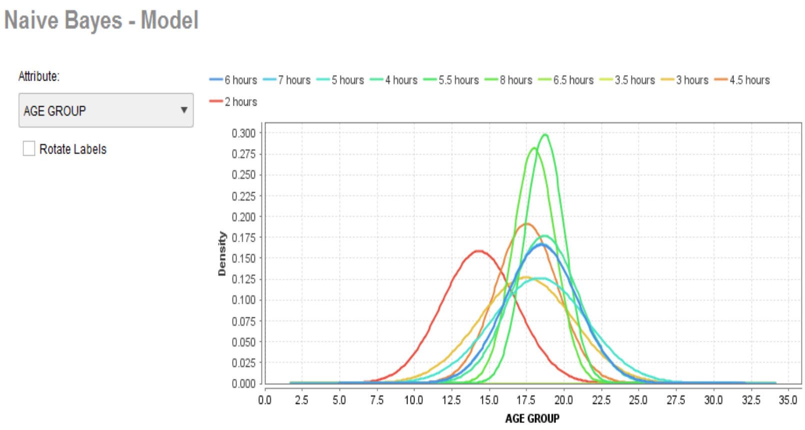
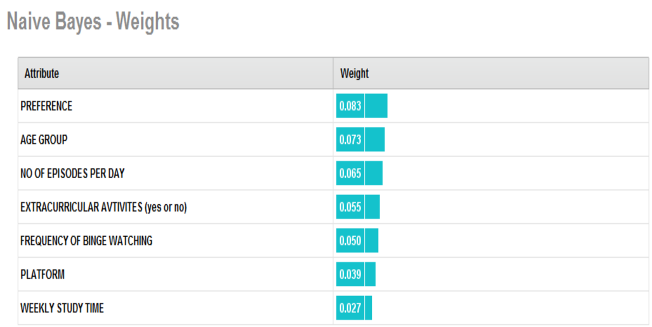


Figure 4: Average duration sleep of youngsters

Naïve Bayes– weights:

According to Naïve Bayes model the attributes are provided with various weightage according to their performance in mining technique. Here, age group provides the accuracy of approx. value 0.073 to this study of binge watching, where as attribute such as platform contributes 0.039 to the study. It provides very good accuracy in reality, even though some attributes are dependent of each other.


Naive Bayes - Weights

Attribute	Weight
PREFERENCE	0.083
AGE GROUP	0.073
NO OF EPISODES PER DAY	0.065
EXTRACURRICULAR AVTIVITES (yes or no)	0.055
FREQUENCY OF BINGE WATCHING	0.050
PLATFORM	0.039
WEEKLY STUDY TIME	0.027

Figure 5: Attributes of the dataset and their corresponding Naive Bayes weights.

Naïve Bayes-simulator:

According to the Figure 6, it is derived that with the help of Rapid Miner, that people at the age group of 18-20(including 18) contributes a vast part to binge watching. And also, their participation in extracurricular activities is very less or zero participation. With the help of Naïve Bayes simulator, we could infer that people at their final teenage stage binge watch regularly provided platform like Netflix-series with least of two episodes.

ISSN:0971-1260
Vol-22-Issue-10-November-2019

According to Figure 7, the time spent for sleeping per day is most likely to be 5.5 hours. Average of the datasets concludes that, people at the age group of 18-20, binge watch with duration of sleep 5.5 hours only. Approximately 30% of the people at specified age group sleep nearly 5-6 hours per day, major part of their daily routine is contributed for binge viewing.

In Figure 7(important factors for 5.5 hours), preference supports the maximum for 5.5 hours of sleep whereas platform supports the minimum to 5.5 hours. Extracurricular activities contradict the maximum to duration of sleep, and age group is the factor that contradicts the minimum. To conclude, preference is directly proportional to duration of sleep and an extracurricular activity is indirectly proportional to duration of sleep.

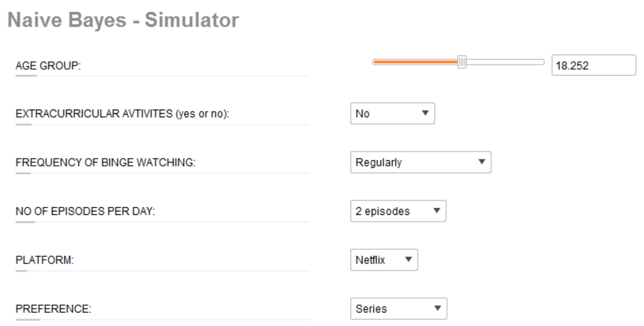


Figure 6: Frequency of dataset attributes for Naive Bayes Classification.

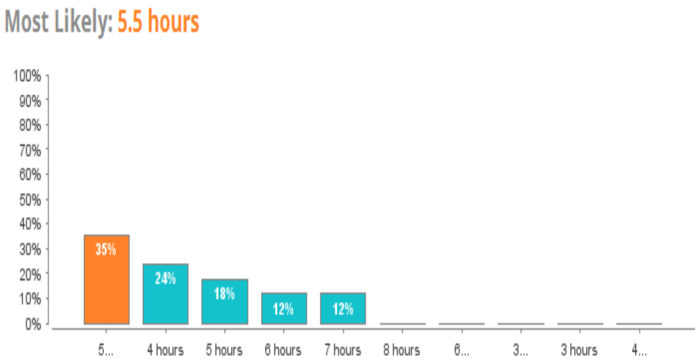


Figure 7: Resultant Bar chart of Naive Bayes Classification.

DECISION TREE ALGORITHM

Decision tree classification:

A decision tree is a branching method which depicts every possible outcome. It's like flowchart structure. The main nodes in decision tree are decision nodes, chance nodes, end nodes. The decision tree is easy to interpret and analyze. The analysis of data is easier in this method as it we get to know every possible outcome. Decision tree was obtained to visualize the association between the structural behaviors and affected states according to training specimen [19].

Decision tree-Weights:

According to decision tree model , the attributes are provided with various weightage according to their performance in mining technique. The frequency of binge watchers provides an accurate value of 0.029 weight ,thus conveying that 29% of the weight of persons are frequent binge watchers. This method provides accurate value for analysis.

ISSN:0971-1260
Vol-22-Issue-10-November-2019

Decision Tree - Weights

Attribute	Weight
PLATFORM	0.065
AGE GROUP	0.056
NO OF EPISODES PER DAY	0.047
WEEKLY STUDY TIME	0.041
PREFERENCE	0.031
FREQUENCY OF BINGE WATCHING	0.029
EXTRACURRICULAR AVTIVITES (yes or no)	0.022

Figure 8: Attributes of the dataset and their corresponding Decision tree weights.

Decision tree-simulator:

From the decision tree simulator, Figure 9, It is inferred that the persons between the age group of 18-25 are frequent binge watchers. The frequency of binge watching is regularly, which means most of people do binge watching regularly as a habit. The frequent platform used is Netflix and the frequently viewed content is series.

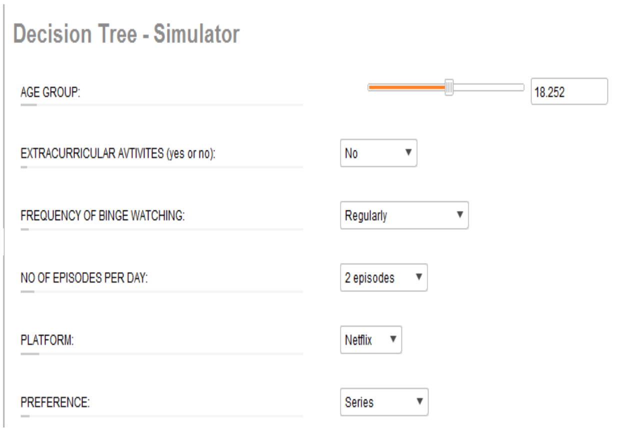


Figure 9 : Frequency of dataset attributes for Decision Tree algorithm.

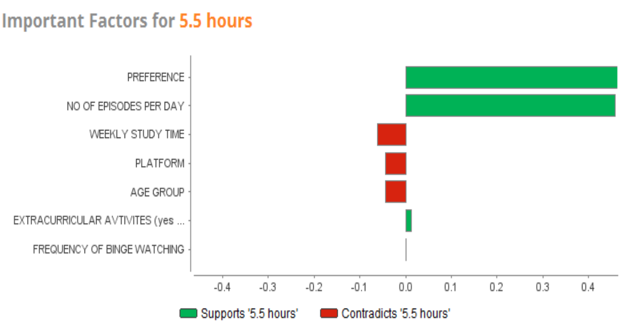


Figure 10: Resultant chart for Decision tree algorithm

Decision tree analysis:

From the decision tree figure 10, It is inferred that the frequently used platforms are Hulu and Netflix wherein the Hulu platform is used for 7 hours and Netflix is used for 6 hours. People prefer watching movies for 4 hours, reality shows for 4 hours and series for 5.5 hours. The average number of episodes per day is 3-5 episodes. The frequency of binge watching is once a week, people binge for 6 hours, regularly people do for 5 hours and twice a week people do for 6 hours and rarely for 5 hours. Hence the inference from decision

ISSN:0971-1260
Vol-22-Issue-10-November-2019

tree is that series is more preferred for watching in the Hulu platform with 3-5 episodes per day and the frequent binge watching is done twice a week.

SUPPORT VECTOR MACHINE ALGORITHM

Support vector machine :

SVM are associated with learning algorithms that analyze data. It's also known as support vector network. They are used for classification and regression purpose. It performs linear calculation as well as nonlinear calculation called The Kernel trick[20]. Classification and mining image data and text data can be easily performed through SVM.

Support Vector Machine - Model

```
Total number of Support Vectors: 119
Bias (offset): -7.658

Feature weight calculation only possible for two class learning problems.
Please use the operator SVMWeighting instead.

number of classes: 11
number of support vectors for class 6 hours: 27
number of support vectors for class 5.5 hours: 10
number of support vectors for class 7 hours: 15
number of support vectors for class 6.5 hours: 1
number of support vectors for class 8 hours: 2
number of support vectors for class 4 hours: 22
number of support vectors for class 3.5 hours: 1
number of support vectors for class 3 hours: 11
number of support vectors for class 5 hours: 23
number of support vectors for class 4.5 hours: 4
number of support vectors for class 2 hours: 3
```

Figure 11: SVM model for Binge Watching.

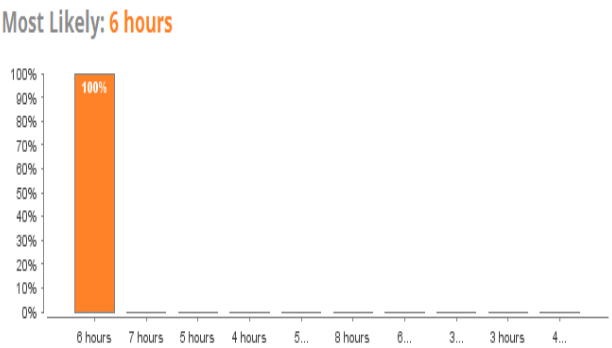
Support vector machine-Weights :

With SVM classification, it is inferred that the frequency of binge watching is 0.075, stating that the binge watching is frequent. The platform has a weight of 0.081 for YouTube ,0.093 for Netflix and 0.064 for Hot star, stating that Netflix is used maximum times. The frequent preferences are movies, which consists a weight of 0.090.

Support Vector Machine - Weights

Attribute	Weight
NO OF EPISODES PER DAY = 2 episodes	0.156
PREFERENCE = Series	0.125
FREQUENCY OF BINGE WATCHING = Rarely during vacations	0.112
PLATFORM = Netflix	0.093
PREFERENCE = Movies	0.090
PLATFORM = Youtube	0.081
FREQUENCY OF BINGE WATCHING = Monthly once	0.075
FREQUENCY OF BINGE WATCHING = Once a week	0.075
EXTRACURRICULAR AVTIVITES (yes or no) = No	0.072
PLATFORM = Hotstar	0.064

Figure 12: Attributes of the dataset and their corresponding SVM weights.



ISSN:0971-1260
Vol-22-Issue-10-November-2019

Figure 13:Average duration of sleep of youngsters with respect to the percentage of binge watching.

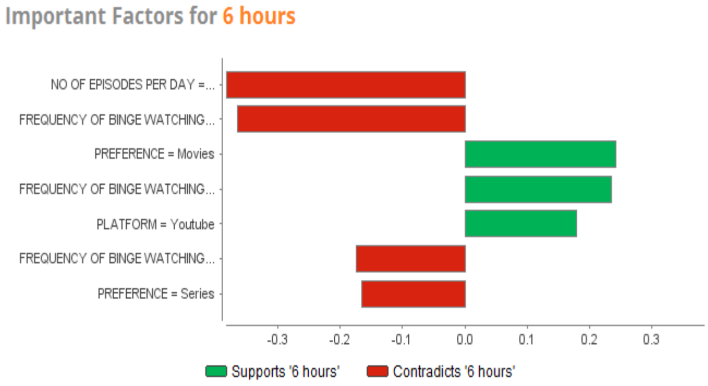


Figure 14: Factors supporting and contradicting sleep duration with respect to Binge watching.

Model	Accuracy	Standard Dev...	Total Time	Training Time (1,0...	Scoring Time (1,0...
Naive Bayes	13.6%	± 4.9%	241 ms	~0 ms	85 ms
Generalized Linear Model	15.5%	± 6.9%	301 ms	1 s	37 ms
Logistic Regression	12.1%	± 4.9%	2 s	378 ms	134 ms
Fast Large Margin	No results yet...	No results yet...	No results yet...	No results yet...	No results yet...
Deep Learning	12.0%	± 4.3%	674 ms	3 s	12 ms
Decision Tree	13.3%	± 9.5%	238 ms	8 ms	12 ms
Random Forest	20.2%	± 9.3%	8 s	25 ms	427 ms
Gradient Boosted Trees	No results yet...	No results yet...	No results yet...	No results yet...	No results yet...
Support Vector Machine	18.8%	± 6.6%	931 ms	42 ms	12 ms

Figure 15: Resultant table of Data Mining algorithms applied to dataset.

With the increase in technology usage influence the increase in the number of platforms, the devices, the programs, the rate of binge also increases. The accuracy from naive based algorithm is 14%, that from the decision tree is 13% and SVM is 19%. Platforms Netflix, Amazon Prime and Hulu are used in maximum numbers. The most viewed content is Series since they are interesting and a variety of genres are present. It is inferred that the age group in which maximum binge watchers is between 15 and 25. We also infer that due to the growth of technology the percentage of binge watching is growing and is expected to grow in abundant number in the future. Along with binge watching, binge eating is also increasing.

VII. CONCLUSION

Unwinding time is significant and for some, binge watching is an outlet to unwind. Avid binge watchers reported poor health quality, weakness and more sleep deprivation symptoms such as insomnia etc. Due to binge watching the attentiveness and briskness in people is decreasing. People tend to lose their sleep and drain out their energy. Due to binge watching the quantity of intake of food is also decreasing in humans. The review depicts, binge watching and poor way of life decisions, for example, choosing unhealthy food, undesirable meals and snacks, sleep deprivation. Technology is affecting human conduct, and how that conduct is influencing individuals' well-being. The high patterns of TV use were interconnected with maximised risk for cardiac disease, low cognitive function in middle of life.

REFERENCES

1. The Netflix Effect: Teens, Binge Watching, and On-Demand Digital Media Trends Sidneyeve Matrix, Published January 15, 2014, The Centre for Research in Young People's Text and Cultures.

ISSN:0971-1260
Vol-22-Issue-10-November-2019

2. Sprinting a media marathon: Uses and gratifications of binge-watching television through Netflix. - Matthew Pittman and Kim Sheehan, Published October 5, 2015, First Monday.

3. Confessions of A 'Guilty' Couch Potato Understanding and Using Context to Optimize Binge-watching Behavior: Dimph de Feijter, Vassilis- Javed Khan, Marnix van Gisbergen, Published June 22, 2016, TVX'16 Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video.

4. New era of TV-Watching behavior: Binge watching and its psychological effects: Azza Abdel-Azim Mohamed Ahmed, Published March, 2019, Media Watch.

5. The Relationships Between Television Viewing Behaviors, Attachment, Loneliness, Depression, and Psychological Well-Being: Katherine S Wheeler, Published April 20, 2015, Other Psychology Commons (UGC).

6. Binge-watching: Video-on-demand, quality TV and mainstreaming fandom - Mareike Jenner, Published September 18, 2015, SAGE Journals.

7.Binge watching and college students: Motivations and Outcomes: Swati Panda, Satyendra C. Pandey, Published 20 November, 2017, Young Consumers.

8. Breaking binge:Exploring the effects of binge watching on television viewer reception : Lesley Lisseth Pena, Published September 29,2015, Social and Behavioral Sciences Commons.

9. Binge viewing, sleep and the role of pre-sleep arousal : Liese Exelmans, Jan Van Bluck, Published August 15, 2017, Journal of Clinic, Sleep and Medicine.

10. Viewing patterns and addictions to television among adults who self-identify as binge watchers : Monita karmakar, Jessica sloan kruger, Jon elhai, Alaina kramer, Published November 2015, American Public Health Association.

11. The Netflix Effect and Defining Binge-Watching : Brenna C Davis, Published April 11,2016, Undergraduate Research Commons.

12. Binge-watching : A suspenseful, emotional, habit : Bridget Rubenking and Cheryl Campanella bracken, Published 03 October, 2018, Communication Reserach Reports

13. Multivariate Relationships of Binge-Watching-Drinking-Eating with depression, anxiety, and stress in college students : Katina Letrice Clarke, Published May 29, 2019, Social Psychology Commons

14. Assessing binge-watching behaviors: development and validation of the “watching TV series motives” and “binge watching engagement and symptoms” questionnaires: Maeva Flavelle and Joel Billieux, Published January 20,2019, Computers in Human Behaviour.

15. On binge-watching : nine critical propositions : Tanya horeck, Mareike Jenner, Tina Kendall,Published December 2018, The International Journal of Television Studies.

16. An exploration of the motivations for binge-watching and the role of individual differences: Hongjin shim and Ki Joon Kim, Published - May 2018, Computers in Human Behavior

17. ‘Just one more episode’: frequency and theoretical correlates of television binge watching: Emily Walton-Pattison, Stephen u Dombrowski, Justin presseau, Published April 22, 2016, Journal of Health Psychology.

ISSN:0971-1260
Vol-22-Issue-10-November-2019

18. Bayesian Naïve Bayes classifiers to text classification: Shuo Xu, Published November 1, 2016, SAGE Journals

19. Decision tree-based seismic damage prediction for reinforcement concrete frame buildings considering structural micro-characteristics: Liang Su, Hai-jian he, Published February 26, 2019, SAGE Journals

20. Kernel parameter selection for support vector machine classification: Zhiliang Liu and Hongbing xu, Published June 1, 2014, SAGE Journals.