COMP8410 ASSIGNMENT 1

Netflix's Recommendation System and its Ethical Issues

Rui Qiu

u6139152

13 March, 2018

BACKGROUND

Data mining does boost the skyrocketing of technologies companies built upon recommendation algorithms. It provides multi-dimensional information towards each service user so that the company can understand its audience and arrange more personalised intending service for them. Nevertheless, some ethical issues arise at the same time. In this essay, we are going to discuss the case of Netflix, its potential ethical concerns, and suggestive fixes.

PROJECT

AIMS

As one of the most competitive streaming services in the market, Netflix could never achieve its accomplishment today without the help of its recommendation system. The direct goal of data mining in this system is to discover the preferences in content viewing; then it can provide personalised streaming experience for different kinds of customers. In this way, Netflix maintains its user stickiness and keeps its market share.

METHODS

So the core of this algorithm is how to find movies or TV plays this specific customer would like to watch. As Netflix's vice president of product innovation, Todd Yellin, mentioned a very interest metaphor in an interview by WIRED last year, "The three legs of this stool would be Netflix members; taggers who understand everything about the content; and our machine learning algorithms that take all of the data and put things together". Everyday, millions of Netflix users generate a huge amount of data such as what people watch, "What we see from those profiles is the following kinds of data – what people watch, what they watch after, what they watch before, what they watched a year ago, what they've watched recently and what time of day" (Plummer, 2017). After that, to understand the raw data better, Netflix has freelancers and in-house staff to tag the content by various standards. And finally, a machine learning algorithm is applied to figure out the behaviour of customers and put some weights over the previous two "legs" and form so-called "taste communities" such that people would be recommended the same kind of things trending among people like them. To conclude the general idea of this algorithm, it is a clustering algorithm done by a mixture of unsupervised and supervised learning. The clustering itself is unsupervised. However, the taggers indeed

provide some extra information as guidance towards the content watched.

SOURCE AND NATURE OF DATA

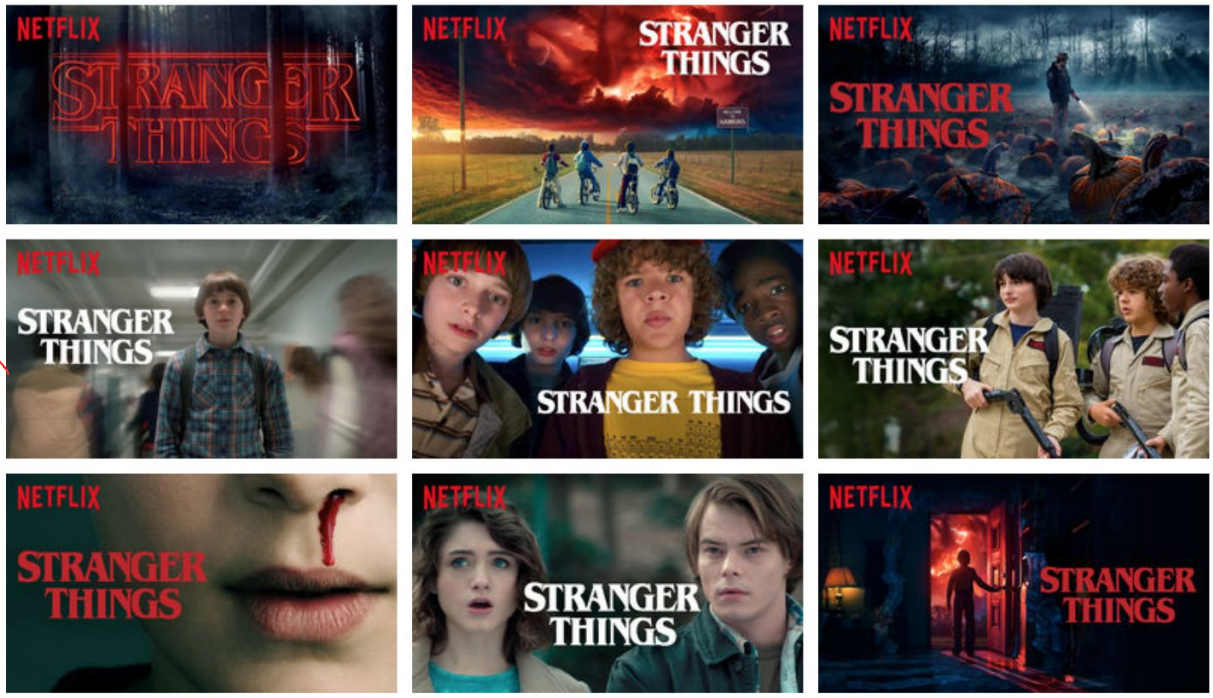
Naturally, a majority of the data comes from what people have watched, whether they finished watching it, etc. But this is not the whole story, the data involved in this mining procedure not only comes from direct feedback, such as a “thumbs-up” after watching or adding to My List but some implicit behaviour data as well. These kinds of data are more “behaviour” related. It is about how you watched a TV series rather than just binary feedback. For example, if you watch it for two nights in a row, instead of few episodes every week, you probably are really into the plot of this series. Moreover, Netflix also collects information about users' playing devices, internet provider, and geo-location, etc.

AUTHORITY OF DATA ACCESS

According to Netflix's Privacy Statement, it uses personal information for the betterment of user experiences but has not ruled out the possibility of disclosing data to third parties. Generally speaking, Netflix retains the right to publish the data if it satisfies legal process or government request. Moreover, the information will be transferred, if Netflix is involved in any reorganisation, restructuring, merger or sale, or other transfer of assets.

EXPECTED AND UNEXPECTED RESULTS

The outcome of this evolving recommendation system cannot be merely unseen since Netflix is, in fact, doing better and better in its streaming media area. Additionally, Netflix has refined its recommendation system with a new feature, which is personalised



visuals based on each member's aesthetics preference (Chandrashekar, Amat, Basilico & Jebara, 2017).

With the support of a great deal of user-provided preferences, Netflix expanded its specialisation into film and television production in 2013. This move is somewhat as expected because as a company understanding the viewers so much, it is a hardly a risky move to make profits on it.

The only thing that overshadows this successful product is the privacy concerns it provokes. In December 2017, a tweet on Netflix's account stating that 53 people have watched A Christmas Prince every day for the past 18 days was considered “creepy” (Murphy, 2017). This tweet also raised a discussion about the privacy of online streaming services. The current of these commercial companies like Netflix and Spotify are starting to use anonymous user data in their marketing campaign (Maheshwari, 2017). But what if the data gathered by Netflix's recommendation system is compromised? What if the algorithm understands too much and backfires? These questions are not currently on the surface, but still unavoidable.

ETHICAL ISSUES

One thing undeniable is that ethical issues are lurking inside the commercial success. In

the following paragraphs, we are going to discuss some potential possibilities.

BIASED DECISIONS

"[Algorithms] replace human processes, but they're not held to the same standards," Cathy O'Neil, the author of Weapons of Math Destruction, says, "People trust them too much." (O'Neil, 2017)

In the context of the recommendation system, we know that basically everything we see within the range of Netflix service is a result of the recommendation. Surely we can trust the algorithm full-heartedly, but it is quite evident that algorithm bias is gradually becoming a social issue along with the development of machine learning. For instance, a customer of a particular ethnicity might be recommended a lot of movies starred by actors or actresses of the same ethnicity; this could have negative consequences for minorities.

Let us take a look at the Principles for Algorithmic Transparency and Accountability issued by Association for Computing Machinery, and try to figure out if biased recommendation results violate these principles.

The first principle states that "owners, designers, builders, users, and other stakeholders of analytic systems should be aware of the possible biases involved in their design, implementation, and use and the potential harm that biases can cause to individuals and society.". That is to say, companies should be aware of the fallout of discrimination inside their recommendation algorithm. Meanwhile, the seventh principle emphasises the role of validation and testing. It claims that algorithms should be tested for discriminatory harm, and make the results public.

However, the status quo of recommendation algorithm is not optimistic, even though no cases about Netflix have appeared, similar issues did happen to large companies like Google, which once confused photos of gorillas and African American people (Simonite, 2018). And this "bug" still haunted them due to the shorthand of current machine learning algorithm. Hence, censoring specific search entries have been the most commonly used solution. Hopefully, the biased view of an algorithm can be resolved more intelligently in the future.

INDIVIDUAL PRIVACY

Another apparent ethical aspect is the privacy of individual information. It includes not only personal information like name, address, payment methods, which are fundamental components of user information as any other online paid services, but also all

behaviour-related information that illustrates an individual. As mentioned in the project description part, Netflix also records how people watch online streaming contents, for a better understanding of its customers.

Storing this vast amount of data seems to agree with the fifth (data provenance) and sixth (auditability) principles of the 7 Principles. In this way, keeping data is useful and necessary. Now here comes an important question, what if these numbers are revealed to the public?

Interestingly, some aggregated statistics of users' behaviours reflect the status of popular culture, and people love to see them as fun facts, even a method to keep themselves on trend.The trending videos on YouTube and year-end summary done by Netflix and Spotify are such examples. But what if these statistics are narrowed down to some specific group of people, even to some anonymous individuals? Sometime, people could feel being targeted or even stalked in this case, that might be the origin of “creepiness” in the Netflix's tweet debate. It is hard to say that people can stay anonymous on Internet today, and personal data leak is always a potential risk. But if we consider such summary statistics as an “aggregated leak”, for what level people can stand it? Therefore, it is should be carefully considered when a commercial company wants to reveal some fun facts about its customers. Otherwise, people might not be entertained as expected.

QUALITY OF LIFE

Streaming services like Netflix do improve the quality of life of their users, but the that does not mean to make all the choices with an excuse like “for their own good”. Among all the data Netflix collects from its users, which include details about internet providers. In 2017, there were reports about some Internet providers like Verizon, AT&T and Telstra were throttling the download speed of Netflix when the data usage is close to the monthly limit (Brodkin, 2017). Admittedly, this action seems reasonable, and it even can be considered as a benevolent move. Still, Netflix does deceive its customers to some extent since it never mentions controlling downloading speed subjectively. 9 In its Terms of Service, it states as follows:

"The quality of the display of the streaming movies & TV shows may vary from computer to computer, and device to device, and may be affected by a variety of factors, such as your location, the bandwidth available through and/or speed of your Internet connection."

That never mentions "by carrier". Hence we believe such action is not suitable. Referring to Australian Computer Society's Code of Professional Conduct, Netflix appears to respect

the enhancement of quality of life, but on the other hand, it violates the code of honesty. Admittedly, Netflix did not bring any cost to its customer but saved internet fares for them potentially. However, honesty is an absolute standard and cannot be compromised. Netflix failed to make the mechanism of net speed throttling public to its users, in the long run, it could cause a problem of distrust.

IMPROVEMENT ON ETHICAL ISSUES

ON BIASED DECISIONS

One possible way to prevent bias from happening is to remove those variables that could cause bias such as ethnicity or to find a way to compensate it. Tan, Caruana, Hooker and Lou implemented a transparent model distillation approach to audit black-box models so that we can hopefully eliminate the impacts of bias.

The other solution is to strengthen the human factor in decision making, for example, partially replace recommendation algorithm with a human curator or combine these two parts. Currently, the algorithm has not achieved a level that makes everyone satisfied so that we could use some help from ourselves.

ON INDIVIDUAL PRIVACY

If we take a look at the history, the overlapping area of APEC[1] Privacy Framework (2005) and OECD[2] Privacy Guidelines (2013), though not entirely similar, can be served as a suitable starting point to establish an outline for the privacy protection on a large scale. Some aspects are highlighted in both:

- **Collection limit.** The bottom line is to reduce the amount of data needed. Not everything is useful, even if some are correlated. Netflix should try to use as little data as possible to generate a recommendation for a user.
- **Specified purpose.** Users should be fully informed the purpose of data collecting before the mining process.
- **Use limitation.** Collected data should not be revealed or used for other purposes unless with the consent of users or by law.

Two branches of fixes should be paid close attention to, one is to use more advanced technical methods, the other is to use the power of legislation.

[1] Asia-Pacific Economic Cooperation.
[2] The Organisation for Economic Co-operation and Development.

Technically, the ultimate method to protect private information like watching habits is data anonymisation. This can be carried out by combining similar data and process with bulk encryption. Especially, for some very similar data. Some other mining methods that preserve privacy include:

1. Randomization methods: Add noise to the data to conceal the raw data.
2. The k-anonymity and l-diversity methods: The k-anonymity makes a given record mapped to at least k other records in the data, and the l-diversity enforces intragroup diversity of sensitive values to ensure anonymisation.
3. Distributed privacy preservation: Partition large dataset by subsets (horizontally) or by attributes (vertically).
4. Downgrading the effectiveness of data mining results: Can hide some data partially or slightly distort some classification rules.
5. Differential privacy: An algorithm which behaves nearly the same on two similar data. (Han, Kamber & Pei, 2011)

Along with the development of technology, the de-anonymisation algorithm is getting more and more powerful. In fact, the second Netflix Prize was suspended due to exposure of customer information to the competitors (Singel, 2010). Overall, technology easily outgrows the pace of legislation these days. This is somewhat not surprising since the legislation usually takes a comparatively long procedure but the burst of technology is in a flash. Therefore, besides technological preventions, we should also strengthen regulatory guidance. Public education about privacy protection is also a significant complement. Only in the way of combining self-discipline of companies and legal supervision, can the right of privacy be guaranteed to individuals. Some specific governance and educational approach can be:

- Accelerate the legislation related to online private privacy.
- Set up supervise commision to ensure that the frontier technology applied to the industry will follow the framework of current regulations, etc.

ON QUALITY OF LIFE

Regarding the discussed ethical issue about the quality of life, companies should be more responsible for the service they provide and be clear about the word usage in their documentations such as terms of service. One truth is pretty evident that, if a company improves the quality of life of its users, it will win a good portion of market share. Otherwise, it will suffer from a loss in the long run. Hence, we could say, a company should target the improvement of human lives for both the greater good and their

interest. Without a doubt, no company would refuse profits from its point of view. There is nothing special we need to emphasise in this aspect.

CONCLUSION

The power conferred on the recommendation algorithm by data mining, indeed changes the industry of online streaming, even the entertainment industry. But it is also a double-edged sword. The risk of privacy being invaded has never been this high in the past. It is a cliché, but saying "with greater power comes with greater responsibility" is never an empty talk. Companies like Netflix should be more responsible for data they collect, to its customers, and by ethical standards. After all, the goal of technological progress we make is to enhance the world, and it should not be built on the cost of breaking any ethical standards.

REFERENCES

- O'Neil, C. (2017) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Broadway Books.
- Han, J., Kamber, M. & Pei, J. (2011) Data Mining: Concepts and Techniques, Third Edition. Morgan Kaufmann.
- Netflix. (2017) *Netflix Terms of Use*. Netflix. Available from: https://help.netflix.com/en/legal/termsofuse [Accessed: 13 March 2018].
- Netflix. (2017) *Netflix Privacy Statement*. Netflix. Available from: https://help.netflix.com/legal/privacy [Accessed: 13 March 2018].
- Graham, R. & Low, G. (2014) *Australian Computer Society Code of Professional Conduct*. Australian Computer Society. Available from: https://www.acs.org.au/content/dam/acs/rules-and-regulations/Code-of-Professional-Conduct_v2.1.pdf [Accessed: 13 March 2018].
- Association for Computing Machinery US Public Policy Council. (2017) *Statement on Algorithmic Transparency and Accountability*. Available from: https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf [Accessed: 13 March 2018].
- CTI Sub-Fora & Industry Dialogues Groups, Electronic Commerce Steering Group. (2005) *APEC Privacy Framework*. Available from: https://www.apec.org/Publications/2005/12/APEC-Privacy-Framework [Accessed: 13 March 2018].
- OECD. (2013) *OECD Privacy Guidelines*. Available from: http://www.oecd.org/sti/ieconomy/privacy-guidelines.htm [Accessed: 13 March 2018].
- Wikipedia. (2018) *Netflix*. Wikimedia Foundation Inc. Available from: https://en.wikipedia.org/wiki/Netflix [Accessed: 13 March 2018].
- Singel, R. (2010) Netflix Cancels Recommendation Contest After Privacy Lawsuit. WIRED. Available from: https://www.wired.com/2010/03/netflix-cancels-contest/ [Accessed: 13th March 2018]

- Narayanan, A. & Shmatikov, V. (2007) *How To Break Anonymity of the Netflix Prize Dataset*. arXiv. [Preprint] Available from: https://arxiv.org/abs/cs/0610105 [Accessed: 13th March 2018].
- Schneier, B. (2007) *Why 'Anonymous' Data Sometimes Isn't.* WIRED. Available from: https://www.wired.com/2007/12/why-anonymous-data-sometimes-isnt/. [Accessed: 13 March 2018].
- Amatriain, X. & Basilico, J. (2012) *Netflix Recommendations: Beyond the 5 stars (Part 1)*. Medium. Available from: https://medium.com/netflix-techblog/netflix-recommendations-beyond-the-5-stars-part-1-55838468f429 [Accessed: 13 March 2018].
- Amatriain, X. & Basilico, J. (2012) *Netflix Recommendations: Beyond the 5 stars (Part 2)*. Medium. Available from: https://medium.com/netflix-techblog/netflix-recommendations-beyond-the-5-stars-part-2-d9b96aa399f5 [Accessed: 13 March 2018].
- Vanderbilt, T. (2013) *The Science Behind the Netflix Algorithms That Decide What You'll Watch Next*. WIRED. Available from: https://www.wired.com/2013/08/qq_netflix-algorithm/ [Accessed: 13 March 2018].
- Simon, P. (2014) *Big Data Lessons From Netflix.* WIRED. Available from: https://www.wired.com/insights/2014/03/big-data-lessons-netflix/ [Accessed: 13 March 2018].
- Govind, N. (2014) *Optimizing the Netflix Streaming Experience with Data Science.* Medium. Available from: https://medium.com/netflix-techblog/optimizing-the-netflix-streaming-experience-with-data-science-725f04c3e834 [Accessed: 13 March 2018].
- Murphy, M. (2017) 'Creepy' Netflix tweet on viewers' TV habits provokes privacy concerns. The Telegraph. Available from: https://www.telegraph.co.uk/technology/2017/12/12/creepy-netflix-tweet-viewers-tv-habits-provokes-privacy-concerns/ [Accessed: 13 March 2018].
- Amatriain, X. (2014) *KDD 2014 Tutorial - the Recommender Problem Revisited.* LinkedIn. https://www.slideshare.net/xamat/kdd-2014-tutorial-the-recommender-problem-revisited [Accessed: 13 March 2018].

- Netflix Technology Blog (2015) *Netflix's Viewing Data: How We Know Where You Are in House of Cards.* Medium. Available from: https://medium.com/netflix-techblog/netflixs-viewing-data-how-we-know-where-you-are-in-house-of-cards-608dd61077da [Accessed: 13 March 2018].
- O'Reilly, L. (2016) *Netflix Lifted the Lid on How the Algorithm that Recommends You Titles to Watch Actually Works.* Business Insider. Available from: https://www.businessinsider.com.au/how-the-netflix-recommendation-algorithm-works-2016-2?r=US&IR=T [Accessed: 13 March 2018].
- Wood, J. (2017) *Originals Born of Data Mining: Netflix's Well-known Secret?* Medium. Available from: https://medium.com/nimbly-aware/originals-born-of-data-mining-netflixs-well-known-secret-6f4f33c011c2 [Accessed: 13 March 2018].
- Plummer, L. (2017) *This is How Netflix's Top-Secret Recommendation System Works.* WIRED. Available from: http://www.wired.co.uk/article/how-do-netflixs-algorithms-work-machine-learning-helps-to-predict-what-viewers-will-like [Accessed: 13 March 2018].
- Kissmetrics Blog. (2017) *How Netflix Uses Analytics To Select Movies, Create Content, and Make Multimillion Dollar Decisions.* Kissmetrics Blog. Available from: https://blog.kissmetrics.com/how-netflix-uses-analytics/ [Accessed: 13 March 2018].
- Netflix Technology Blog (2016) *Selecting the best artwork for videos through A/B testing.* Medium. Available from: https://medium.com/netflix-techblog/selecting-the-best-artwork-for-videos-through-a-b-testing-f6155c4595f6 [Accessed: 13 March 2018].
- Chandrashekar, A., Amat, F., Basilico, J. & Jebara, T. (2017) *Artwork Personalization at Netflix.* Medium. Available from: https://medium.com/netflix-techblog/artwork-personalization-c589f074ad76 [Accessed: 13 March 2018].
- Maheshwari, S. (2017) *Netflix and Spotify Ask: Can Data Mining Make for Cute Ads?* The New York Times. Available from: https://www.nytimes.com/2017/12/17/business/media/netflix-spotify-marketing.html [Accessed: 13 March 2018].

- Van Otterlo, M. (2018) *Gatekeeping Algorithms with Human Ethical Bias: The ethics of algorithms in archives, libraries and society*. arXiv. [Preprint] Available from: https://arxiv.org/abs/1801.01705 [Accessed: 13 March 2018].
- Schneck, A. (2018) *The Subliminal Trick Netflix Uses to Get You to Watch Its Movies & Shows.* Thrillist. Available from: https://www.thrillist.com/entertainment/nation/how-new-netflix-recommendation-algorithm-works [Accessed: 13 March 2018].
- Paraschakis, D. (2016) *Recommender Systems from an Industrial and Ethical Perspective.* ACM Digital Library. Available from: https://dl.acm.org/citation.cfm?id=2959101 [Accessed: 13 March 2018].
- Netflix Media Center (2017) *2017 on Netflix - A Year in Bingeing*. Netflix Media Center. Available from: https://media.netflix.com/en/press-releases/2017-on-netflix-a-year-in-bingeing [Accessed: 13 March 2018].
- Szóka, B. (2016) *Beyond Netflix's Hypocrisy: The Real 'Throttling' Debate.* Medium. Available from: https://techpolicycorner.org/beyond-netflix-s-hypocrisy-the-real-throttling-debate-900c46c6ba79 [Accessed: 13 March 2018].
- Bhargava, R. (2018) *The Algorithms Aren't Biased, We Are.* Medium. Available from: https://medium.com/mit-media-lab/the-algorithms-arent-biased-we-are-a691f5f6f6f2 [Accessed: 13 March 2018].
- Miller, C. C. (2015) *When Algorithms Discriminate.* The New York Times. Available from: https://www.nytimes.com/2015/07/10/upshot/when-algorithms-discriminate.html [Accessed: 13 March 2018].
- Knight, W. (2017) *Biased Algorithms Are Everywhere, and No One Seems to Care.* MIT Technology Review. Available from: https://www.technologyreview.com/s/608248/biased-algorithms-are-everywhere-and-no-one-seems-to-care/ [Accessed: 13 March 2018].
- Simonite, T. (2018) *When it Comes to Gorillas, Google Photos Remains Blind.* WIRED. Available from:

https://www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/ [Accessed: 13 March 2018].

- Brodkin, J. (2017) *Verizon accused of throttling Netflix and YouTube, admits to "video optimization".* Ars Technica. Available from: https://arstechnica.com/information-technology/2017/07/verizon-wireless-apparently-throttles-streaming-video-to-10mbps/ [Accessed: 13 March 2018].

- Snow, J. (2017) *New Research Aims to Solve the Problem of AI Bias in "Black Box" Algorithms.* MIT Technology Review. Available from: https://www.technologyreview.com/s/609338/new-research-aims-to-solve-the-problem-of-ai-bias-in-black-box-algorithms/ [Accessed: 13 March 2018].

- Netflix Technology Blog (2016) *Protecting Netflix Viewing Privacy at Scale.* Medium. Available from: https://medium.com/netflix-techblog/protecting-netflix-viewing-privacy-at-scale-39c675d88f45 [Accessed: 13 March 2018].

- Tan, S., Caruana, R., Hooker, G. & Lou, Y. (2018) *Auditing Black-Box Models Using Transparent Model Distillation With Side Information.* arXiv. [Preprint] Available from: https://arxiv.org/abs/1710.06169 [Accessed: 13 March 2018].

- Reynolds, M. (2017) *Bias test to prevent algorithms discriminating unfairly*. New Scientist. Available from: https://www.newscientist.com/article/mg23431195-300-bias-test-to-prevent-algorithms-discriminating-unfairly/ [Accessed: 13 March 2018].