Neural Network-Based Recommender Algorithms
(Technical Paper)

The Bias of Recommender Systems and Impact on Social Culture

(STS Paper)

A Thesis Prospectus Submitted to the

Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree

Bachelor of Science, School of Engineering

Richard J Park

Fall, 2020

Technical Project Team Members

Hongning Wang

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Signature ____ Date ___________

Richard Park

Approved ____ Date ___________

Aaron Bloomfield, Department of Computer Science

Approved ____ Date ___________

Sean Ferguson, Department of Engineering and Society

Introduction

Recommender Algorithms provide quality results that are critical to the success of online services like YouTube, Amazon, Instagram, and Spotify, but are viewed as a black-box that provides results. Many people are suspicious or skeptical when it comes to how an algorithm came up with a preference, therefore, the need for explainable recommendations is growing simultaneously with the growth of e-commerce and digital viewership. With the introduction of the Netflix Prize competition, a public competition run by Netflix that rewards the team that can beat the company's current movie recommendation system, it is evident that some techniques are superior to others in recommender systems. Through my technical capstone research, I will be implementing and analyzing recommender algorithms by applying Recurrent Neural Networks (RNN) to this domain of recommendations. In a broad sense, most recommender algorithms have similar strategies that use either content filtering or past user behavior. This capstone project will combine existing solutions of RNNs and content filtering techniques to study the implementation of new recommendation systems as well as the evaluation of explainable recommendations. The goal of this project is to tie this into the STS research and uncover the black-box of neural network recommendations into explainable recommendations.

A successful recommendation system implies that there is positive customer feedback, but also poses the question of whether the algorithm that produced the first recommendation is responsible for the creation of an endless recursive loop of recommendations. In an era where content-based recommendations are crucial to online services, it is nearly inevitable that some user will select their recommended product from an algorithm. Once this recommendation has been successfully clicked, our preferences that were once used for the algorithm that originally helped provide recommendations is now building the preferences that we call our own taste.

Interacting with these recommender systems brings ethical challenges. The recommender algorithm acts as an agent that impacts the shaping of individual preferences by serving as a guide to the choices, we make online. The goal of this STS research is to explore the agents involved in creating our digital sense of taste. The STS research will focus on the changes brought to consumer culture from the algorithm's recommendation.

Technical Topic

One of the key aspects of the technical research is to evaluate the current existing implementations of recommender systems in order to gain a knowledge of the black-box nature of these systems and provide an explainable recommendation. An explainable recommendation is one where a recommendation is not given blindly, but rather through provided explanations of the system and how it came to this conclusion. This is quickly gaining more attention as the user base for online services grow, but the recommender systems are still black-boxed. By exploring the implementation of recommendation systems, I will be able to gain a deep understanding or how algorithms are able to make theses recommendations, thus allowing an explainable recommendation.

Personalized recommendations are based on two main strategies of content filtering and collaborative filtering. Content filtering involves a method of creating profiles for each user that contains characteristics to classify the nature of the user. Content filtering requires the need for external information that may not always be available, but can be gained by questionnaires, etc. A successful example of this is the Music Genome Project used by the internet radio service, Pandora (Koren et al.,2009).

With content filtering, there are more predictions based on attributes and characteristics of the profile of the user, while something like collaborative filtering relies on previous behavior without creating an explicit profile for the user based on characteristics. Collaborative filtering is a recommender approach that takes advantage of using past consumer behavior like ratings or transactions in order to produce the profile for each user's preferences. This may be an issue when not a lot of past consumer behavior data is available, therefore, this approach seems less superior than content filtering when dealing with new products and new users. However, when data is available, this approach can utilize latent factor models like matrix factorization. In short, matrix factorization is a successful recommendation technique that uses matrices for input data with one dimension representing users, and the other representing items of interest. A high correspondence between the items of interest and the user results in a recommendation, but this approach is difficult to use when recommendations must be based on short-session data (Hidasi et al., 2015).

The results from the algorithm may be measured through statistical error measures like RMSE, but online services measure the results from recommender systems based on click/view-through rates and user satisfaction in production. This poses the question of whether metrics for the recommender systems are defined by the engineers or the business model of the company. Throughout this project, I will also be working to find the correlation between these metrics as they are both vital for the goal of a successful recommender system.

I will be working with Assistant Professor Hongning Wang in the Department of Computer Science at the University of Virginia, Renqin Cai, and the CS department lab team. The lab has already dedicated efforts to achieve quality recommendation results using neural network models, and I will be continuing to help the efforts of using cutting-edge technology to potentially develop a new recommendation model. Though a new implementation is not an end goal of this project, the goal is to achieve evaluations for explainable recommendations. This project involves using real-world datasets and testing of the algorithms with datasets from Amazon and Spotify.

STS Topic

Recommendation algorithms provide an easier decision-making process to a user's digital experience. It also provides possible suggestions for users to explore in order to match their taste. It is the goal of an online service to provide a consumer with the most relevant and matching product, but since the data used by the algorithms are protected by privacy laws, the consumers are not aware of how their recommendations are created and thus fosters a relationship between users and recommender systems that can infringe on the autonomy of users. It is important to know how recommendations are produced in order to fully understand why we receive them and to be aware of the possible biases presented when we are recommended something by an algorithm. They are important for our digital experience and are necessary for the success of online services, but these recommendation systems hold such power to shape the ways in which audiences discover music, movies, and news. With this power, the recommendation algorithms are designed to act as an agent in computationally shaping popular culture (Morris, 2015).

The rapid growth in digital commodities and services further brings the need for recommender systems to predict user preferences, but the array of algorithms from song suggestions to movie recommendations monitors our tastes according to the corporate logics. By using the actor network theory as a framework for analysis, we can identify the service providers, recommender systems, and user all as agents that affect the culture of our society. By first observing the service providers like Netflix or Spotify, their use of content-based or collaborative filtering approaches poses ethical challenges.

When user profiles are created during the collaborative filtering algorithms, they may unintentionally produce biases as a side effect of creating profiles out of data collected by user activity. This concept is referred to as a type of algorithmic profiling (Milano et al., 2020). When an algorithm constructs models of users that reproduce social categories, they unintentionally introduce biases if the produced social categories do not align with our recognizable social categories. Our online identities are reflected by the algorithmic categorization and promote the idea of personal identity that is dynamically changing with every action we take online. Since the recommender system's model is continuously changing, this type of labelling often does not match with our self-identifying labels. The recommender systems may consider different attributes more significant than others simply based on the goal of the services. Our goal as users are much different than the corporate service, therefore it is likely that the suggestions given are not a reflection of ourself, but rather a suggestion that would most benefit the recommender system and service provider.

The interest of online services is to make their items available and have recommendations that lure consumers to their products. The goal of the recommender system is to make the recommendations based on the given input data and other metadata that the company may have access to. The recommendation is then presented to the users, who act upon these recommendations and have an interest in receiving the most relevant recommendations. Once the recommendation is received, there exists a feedback loop in content-based recommendations where the user can interact back with the system to provide better recommendations in the future.

This feature of recommendations poses ethical questions of manipulation that can ultimately affect our culture.

Recommender systems may appear as “sticky traps” in which their purpose is to entice and hook user into long time usage of their services (Seaver, 2018). In the long term, this causes biases in the recommender systems that encroach on the autonomy of users. By providing explainable recommendations, this helps guard against these biases and help users make decisions that allow them to use recommendations as aids instead of traps. Simultaneously, certain recommendations may generate a self-reinforcing pattern in which the recommended item will continue to be recommended if it was successfully identified as a good suggestion amongst other users.

Recommender systems are vulnerable to unforeseen groups whose goal is to manipulate the feedback cycle between the user and system (Milano et al., 2020). For example, if a group of active users were to interact with the recommender system and drive up positive feedback for certain items/services, it is likely that the item will be recommended for others. This can be the case for social networks, streaming platforms, and news systems. The nature of content-based filtering isolates users into a bubble of self-reinforcing ideologies that limits exposure to contrasting viewpoints since these contrasting ideas do not result in more user retention. This social effect is damaging to society and harms the function of public debate and democratic institutions (Milano et al., 2020). By this method, recommender systems are vulnerable to propaganda attacks in the circulation of media, ultimately effecting the ways in which information is presented to us.

The stakeholders of recommender systems involved the services, systems, and users. All of these stakeholders play a role as actants in shaping the algorithmic culture. Since most data is protected, research will be achieved through the implementations of different recommender systems, specifically focusing on the effects of collaborative filtering, content-based strategies, and neural network improvements. Many developers are working to avoid the issues mentioned, specifically, to prevent groups of active users from manipulating the recommender systems by performing manipulated feedback cycles.

Next Steps

Explainable recommendations can give consumers comfortability in trusting a recommender system by taking actions that are not manipulated by algorithmic suggestion. If users are aware of certain biases that may be present, users will overall have a more positive impact on network of the feedback cycle. Neural Networks are cutting edge technology that has been proven to improve recommendations. Throughout this research project, I aim to gain experience with this technology and hope to identify the proper metrics of a good recommendation that minimizes bias and provides good suggestions rather than manipulation.

I will also explore the ethical questions behind recommender systems such as issues of user privacy. During the Netflix Prize competition, researchers were able to gather enough information to make inferences about “anonymous” users and were able to individually identify people based only on movie rating preferences. To protect user privacy, recommender systems must work in such a way to keep user identity private. Moving forward these are some of the questions I would like to answer so that recommendations can have a meaningful and positive impact on a technology dependent society.

References

Baltrunas, L., Kaminskas, M., Ludwig, B., Moling, O., Ricci, F., Aydin, A., Lüke, K.-H., & Schwaiger, R. (2011). InCarMusic: Context-Aware Music Recommendations in a Car. In C. Huemer & T. Setzer (Eds.), *E-Commerce and Web Technologies* (pp. 89–100). Springer Berlin Heidelberg.

Beer, D. (2009). Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society*, *11*(6), 985–1002. https://doi.org/10.1177/1461444809336551

Cotter, K. (2019). Playing the visibility game: How digital influencers and algorithms negotiate influence on Instagram. *New Media & Society*, *21*(4), 895–913. https://doi.org/10.1177/1461444818815684

Hallinan, B., & Striphas, T. (2016). Recommended for you: The Netflix Prize and the production of algorithmic culture. New Media & Society, 18(1), 117–137. https://doi.org/10.1177/1461444814538646

Hargittai, E. (2007). The Social, Political, Economic, and Cultural Dimensions of Search Engines: An Introduction. *Journal of Computer-Mediated Communication*, *12*(3), 769–777. https://doi.org/10.1111/j.1083-6101.2007.00349.x

Maguire, J., & Matthews, J. (2012). Are we all Cultural Intermediaries Now? *European Journal of Cultural Studies*. https://doi.org/10.1177/1367549412445762

Morris, J. W. (2015). Curation by code: Infomediaries and the data mining of taste. *European Journal of Cultural Studies*, *18*(4–5), 446–463. https://doi.org/10.1177/1367549415577387

Seaver, N. (2019). Captivating algorithms: Recommender systems as traps. *Journal of Material Culture*, *24*(4), 421–436. https://doi.org/10.1177/1359183518820366

Y. Koren, R. Bell and C. Volinsky, "Matrix Factorization Techniques for Recommender Systems," in *Computer*, vol. 42, no. 8, pp. 30-37, Aug. 2009, https://doi.org/10.1109/MC.2009.263