Personalizing the Digital Experience: Using Machine Learning Models for an Appealing User Experience

Politics of Designing a Digitally Personalized Experience

A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

By

Kaleb Getachew

November 1, 2021

Technical Team Members: John Olamofe, Chang Xu

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.

ADVISORS

Sean Ferguson, Department of Engineering and Society

Nitin Kishen, Software Engineering Manager at PayPal

Introduction

Within a daily occurrence, people are faced with decision-making situations countless times, many of which are unnoticed. Naturally, experts or friends are the first point of contact to aid in decision making, but with the widespread use of the internet of things within the 21st century, the transfer of information has become very rapid, leading to the tracking and storing of user data for data driven personalization.

Throughout the summer of 2021, my responsibilities as a Software Engineering intern at PayPal included all things related to preparation of the company's suite for world markets. I worked in conjunction with two other interns and various other professionals from different departments to ensure users from all sectors were tended to and had a meaningful experience while interacting with the company's suite. Furthermore, the team worked specifically on converting, preparing, and personalizing Polish data for live use this upcoming year. The overarching task was to improve user satisfaction whenever users would visit the page, through the use of relevant natural language processing (NLP) models. This meant ensuring everything displayed is easily understandable, accessible, and most of all relatable for each user. All of which was achieved based on data collected on each individual user.

At the core of creating a personalized digital experience, recommender systems play a huge role in engaging users. With the recent advances in AI, data analytics and big data over the past years, opportunities have arisen for recommender systems to embrace the impressive achievements of AI. Recommender systems create advanced insights into the relationships between users and items, presenting more complex data representations, and discovering comprehensive knowledge in demographical, textural, virtual and contextual data. Big Tech

Companies (BTCs) like Netflix create a personalized web/digital-experience based on consumer collected data to intentionally improve overall customer satisfaction amongst other unintended consequences. Throughout the remainder of the paper, the Oxford Academic definition of personalization will be used; it is the “strategic creation, modification, and adaptation of content and distribution to optimize the fit with personal characteristics, interests, preferences, communication styles and behaviors”(Bol, 2018). This paper will aim to investigate the effect of data collection and data usage by BTCs in the creation of an imaginary personalized digital experience.

Technical Topic

At the start of the internship, my team was tasked with increasing the overall quality effect of PayPal from 77%, based on a survey conducted April of 2021. This metric covers many facets that factor into creating a meaningful and personalized experience for users, ranging from the actual content, to the color scheme, user interface and overall ease of access and interpretability. Other companies within the same industry had percentages well above 90% which meant that PayPal was not able to provide an engaging user experience in comparison to their competitors. Given a website, we were tasked to personalize the website’s content in such a way the users felt accommodated and welcomed. The team was tasked with specifically starting to improve the quality effect within a small sector before applying the same concept to a much larger market audience. With several managers being fluent in Polish the team decided to begin personalizing the PayPal Polish websites. As a result, the goal of the team was to build the infrastructure, and language models for the creation of recommender systems; below is a high-level pipeline of the project.

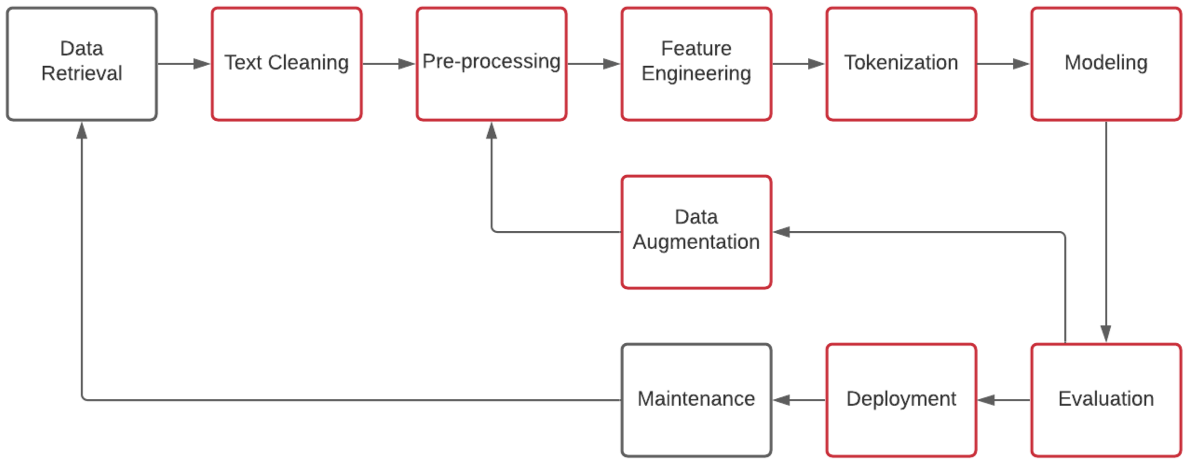

Figure 1. High-level NLP Model Training Pipeline for PayPal's Polish market

The entire pipeline was built for persistence even after the current work that the team has completed. The events with a red-border are ones that my team was directly responsible for; these can be split into four main stages. Stage 0 consists of corpus/data-retrieval, text-cleaning, pre-processing and feature engineering. The corpus is first retrieved from the PayPal servers using multiple SQL queries. The corpus is a mix between PayPal data, which originates from the user's self-disclosed information as well as Polish parliamentary data taken from a public repository. In total, the entire dataset was 572MB in size which is approximately 286,000 pages of textual data. This data is then pre-processed and cleaned before training the language models. We focused on padding any Polish pronouns that were found within the text, so as to avoid any bias or unintended relations; as well as removing any numbers or alpha-numeric characters. These are the preliminary steps that are needed prior to any sort of introduction to machine learning models. Stage 1 consists of tokenization and modeling. The core aspect of the entire goal relied on a text-alteration/generation language model that is able to understand words, and make inferences based on context and various other features. We decided to utilize BERT (Bidirectional Encoder Representations from Transformers), a language model created by researchers at Google AI language. It is innovative in the sense that it applies a bidirectional training of Transformers, as a result it is able to gain a deeper understanding of language context and flow than a typical single-directional language model (Horev 2018). Tokenization and modeling are the most strenuous and time-consuming stages. After the data has been pre-processed we need to be able to tokenize the entire cleaned dataset for the language models to be able to interpret and give weight to the various words; we do this via tokenization. We use BERT's very own tokenizer to feed into the BERT language model. Due to the nature of the dataset, the language model took nearly 20 days to complete its training, using a 20-80 training, testing split. However, to make better use of time, we had decided to run the model on a stratified sample of data to see the effects and results of the language model and apply changes as necessary. Stage 2 consists of evaluation of the model. For this we measured various metrics such as the F1-score, model accuracy, and AUC. Depending on the outcome of the metrics and our satisfaction we can either attempt to improve the model through data-augmentation (creates robust data) or proceed with deployment of the model. We utilized a number of data-augmentation techniques such as Synonym Replacement, Random Deletion, Random Swap, and Random Insertion. These techniques were able to diversify our current data without the need for searching for additional corpra. Stage 3 consists of deployment. Once we were satisfied with the metrics of our language models, we proceeded to pickle (form of language model deployment) the models and deploy them on to an API for company-wide usage. This was achieved via the MongoDB and Node JS stack, which were PayPal standards. From this step, the language models were then integrated to the pre-existing recommender systems for different types of credit payment plans within PayPal. Finally, the websites were then updated based on the final language models and recommender systems and are currently undergoing maintenance and revisions before being migrated to a larger world market.

STS Topic

In order to explore the effect and perception of digital personalization, we can look at a specific case study in a much larger context of BTCs that utilize user data for the sake of customer satisfaction and retention. One such company is Netflix and their use of recommender systems and algorithms.This STS topic will evaluate the consequences as well as the potential use cases of creating a personalized digital experience through the use of recommender systems.

Netflix was founded in 1997 in the United States. It was originally intended to be a mail video club, wherein users would order movies on their website and receive them via mail. With time, the company launched video-streaming services as a complimentary service to their dvd mailing enterprise. The popularity of the on-demand service grew exponentially (from 2007 to the present) and the streaming was offered as a stand-alone service in November of 2010. Netflix as a BTC has a goal of “entertaining the world, whatever your taste, and no matter where you live, we give you access to best-in-class TV shows, movies, and documentaries. We’re streaming in more than 30 languages and 190 countries, because great stories can come from anywhere and be loved everywhere.” (Netflix, 2021). Netflix imagines a future where people from across the world and languages rely on their platform to be entertained. One of their clear and powerful tools of achieving this futuristic goal has been their recommender systems and algorithms. However, much is entailed to achieve a goal of that magnitude. For example, studies the implications of algorithms on the creative process while crafting Netflix's original title *'House of Cards'*, and the possible outcomes and consequences of those decisions for user's behavior and further content creation. He argues that "this app has assembled a sophisticated algorithm model for describing the cultural relationships among individual film and television works, a model that fully embraces the gap between computation and culture" (Finn, 2017). Coupled with their yearly rise in subscribers, from only having 22 million in 2011, to skyrocketing to 214 million in 2021, their influence seems to only be exponentially increasing. This begs the question of Netflix's moral and ethical responsibility, if any, to moderate what they publish towards a global audience.

The digital streaming industry is very unique from any other industry in the fact that it is able to create an immersive experience, both visually and audibly, for the individual user to learn and experience various stories and tales from across the real and fictional worlds. Throughout history the concept of diversity has evolved and its evolution has been fundamental in shaping relationships between people (Grenier, 1989). Meaning that through diversity people are able to fully understand each other, and their thoughts and ideas. There are three main perspectives when relating to diversity: *individual autonomy, deliberative* and *adversarial (*Helberger et al., 2018). Under an individual autonomy perspective, the idea is to give individuals a tool to exploit their different interests. In this case, we can imagine recommender systems and algorithms helping people in diversifying the film experience, broadening the possible choices with regards to their preferences. Pursuing a deliberative perspective, the aim is to promote the public debate, showing divergent opinions and helping people in constructing a critical view. Here, recommender systems can be designed to make users explore films far from their preferences, to make them aware of the unknown parts of the film panorama. With an adversarial perspective, the focus is to broaden the debate highlighting non-dominant visions. Similar to the previous case, recommender systems can serve as a way to promote underrepresented groups, whether subcultures or non-mainstream film styles, through a non-regulated manner.

More recently, their newest addition '*Squid Game',* has been the most-in-demand original series they have produced (Fischer, 2021). This matters because it could very well be the first non-English Netflix series to top its most-watched list. This allows their users to explore a variety of films from around the world, peaking their individual autonomy. Unlike most binge-released series, particularly Netflix Originals, the global audience demand for *'Squid Game'* has increased every single day since its debut on Netflix on Sept. 17. As of Oct. 2, *'Squid Game'* was 102.9 times more in-demand than the average show worldwide — a 481% increase in global audience demand from its launch day (Fischer, 2021). Long were the days in which a language barrier has separated people from viewing films and shows. This introduction to a foreign film introduces a new culture, language, food and social dynamics that might not have even crossed the user's mind before watching the show.

Furthermore, through the deliberative and adversarial perspectives, Netflix has had a wide array of films and documentaries covering hot topics throughout the recent years. Films like *'the Social Dilemma', 'American Factory', 'Two Distant Strangers'* and many more deal with relevant issues, promoting public debate and representing diverging opinions. Films and documentaries like these have had significant buzz during their release, and have sparked the writing of many articles and blogs regarding the issues they have tackled. In addition, they have also provided small film companies the opportunity to display their work to a larger audience. More importantly, by providing the platform for these films and small film companies they have created a powerful tool to start civil discussions.

From the perspective of minority communities and groups, Netflix seems to be on the positive side. Netflix supports a number of social causes like Pride Month, Black History Month, LatinX, Environmental Sustainability and many more. Netflix is an inclusive BTC that strives for diversity within itself and its content, striving “to make it feel like a home, a community of celebration. For underrepresented groups all around the world, there's a lot of hardships in society, a lot of things that bring you down. So we really wanted it to be a place of joy and representation." (Roettgers, 2020). For many people, social media and films influence people in an unforeseen way, people are defined by what they consume and choose to consume what they want to be defined by (Grenier, 1989). This places a huge amount of power in Netflix’s hand. Creating a digital personalized experience entails many seen and unseen consequences, with even more potential use cases. The limit for social good is only defined by the executives of Netflix.

This study aims to determine the consequences and potential use cases of BTCs, using Netflix as a primary example. By fully understanding the politics that are needed to create a digital personalized experience, users and BTCs alike will be able to take part in being entertained as well as effectively shedding light on important hot social and political topics in the current day.

Next Steps

Although the actual project itself is completed, some potential next steps for the technical topic would be to continue researching various types of recommender systems as well as algorithms that take advantage of heavy user data. Exploring various types of Machine Learning and Artificially Intelligent systems would provide more of an insight into the life cycle of user data and provide an explanation, in context, of what becomes recommended for each user.

In regards to the next steps for the STS topic, I will be continuing to gather scholarship on how various groups of people and communities are affected by Netflix. More specifically, I will look towards minority groups and their representations within Netflix films and how their sentiments and goals are perceived. Additionally, exploring different streaming platforms to compare and contrast with Netflix would seem beneficial. It would be interesting to see how other BTCs, aside from Netflix, are using their user's data to improve, or exploit, the digital experience for their customer base.

References

Biddle, Gibson. “A Brief History of Netflix Personalization (Part One, from 1998 to 2006).” Gibson Biddle's "Ask Gib" Product Newsletter, Gibson Biddle's "Ask Gib" Product Newsletter, 17 May 2021, https://askgib.substack.com/p/a-brief-history-of-netflix-personalization.

Bol, Nadine, et al. “Understanding the Effects of Personalization as a Privacy Calculus: Analyzing Self-Disclosure across Health, News, and Commerce Contexts†.” *Journal of Computer-Mediated Communication*, vol. 23, no. 6, 2018, pp. 370–388., https://doi.org/10.1093/jcmc/zmy020.

Dean, Brian: “Netflix Subscriber and Growth Statistics: How Many People Watch Netflix in 2021?” *Backlinko*, 7 Oct. 2021, https://backlinko.com/netflix-users.

Finn, E. (2017). What algorithms want: Imagination in the age of computing. Cambridge, MA: MIT Press

Fischer, Sara. “‘Squid Game’ Becomes a Massive Hit.” Axios, 5 Oct. 2021, https://www.axios.com/squid-game-netflix-massive-hit-d68c4890-bfc4-4255-bb98-db8b36347617.html.

Grenier, L. (1989). From diversity to difference: The case of socio-cultural studies of music. New Formations, 1989(9).

“History of Recommender Systems: Overview of Information Filtering Solutions.” Onespire Ltd. - SAP and Management Consulting, 20 Sept. 2021, https://www.onespire.net/news/history-of-recommender-systems/.

Horev, Rani. “Bert Explained: State of the Art Language Model for NLP.” *Medium*, Towards Data Science, 17 Nov. 2018, https://towardsdatascience.com/bert-explained-state-of-the-art-language-model-for-nlp-f8b21a9b6270.

Jackson, Dan. "The Netflix Prize: How a $1 Million Contest Changed Binge-Watching Forever." Thrillist, https://www.thrillist.com/entertainment/nation/the-netflix-prize.

Koren, Yehuda. "The BellKor Solution to the Netflix Grand Prize" Netflix Prize, Aug. 2009, https://www.netflixprize.com/assets/GrandPrize2009_BPC_BellKor.pdf.

Matrix, Sidneyeve. (2014). The Netflix Effect: Teens, Binge Watching, and On-Demand Digital Media Trends. Jeunesse: Young People, Texts, Cultures. 6. 119-138. 10.1353/jeu.2014.0002.

Netflix Technology Blog. "Selecting the Best Artwork for Videos through A/B Testing." The Netflix Tech Blog, Medium, 3 May 2016, https://netflixtechblog.com/selecting-the-best-artwork-for-videos-through-a-b-testing-f6155c4595f6.

"Privacy Statement." Netflix Help Center, https://help.netflix.com/legal/privacy#cookies.

Roettgers, Janko. "How Netflix Reinvented Its Marketing on Social Media." Protocol, Protocol - The People, Power and Politics of Tech, 13 Oct. 2020, https://www.protocol.com/netflix-marketing-to-targeted-audiences.

"Stories Move Us.they Make Us Feel More Emotion, See New Perspectives, and Bring Us Closer to Each Other." *About Netflix*, https://about.netflix.com/en.

Zhang, Q., Lu, J. & Jin, Y. Artificial intelligence in recommender systems. Complex Intell. Syst. 7, 439–457 (2021). https://doi.org/10.1007/s40747-020-00212-w.