Visualising User Embedding of an NCF Recommender Model – A Creative Approach on Improving Media and Information Literacy

by
HSI FANG CHEN
20002791

This thesis is submitted in partial fulfillment of the requirement
for the award of the degree:

MSc Data Science and AI for the Creative Industries
Creative Computing Institute, University of the Arts London
Supervisor: Dr Louis McCallum

19 September 2022

Abstract

The paper adopted an interdisciplinary method to explore a creative approach to making the personalisation algorithm approachable to average users. The first section of the paper introduced the background of automated personalisation in machine learning and the concerns it raised amongst the public. In this thesis, I attempted to provide a perspective on empowering the users through the media and information literacy framework. I adopted the Microsoft news dataset with anonymous user data and recreated a scenario for new participants. A collaborative filtering model is used to generate user embeddings. Lastly, with p5.js, numerical embeddings are turned into 3D space so users can explore its structure. Through a user feedback survey, the statistical result indicates that the method proposed in the thesis has successfully informed the users to be more aware and literate of algorithmic news production.

Introduction

Personalisation on News Platforms

The term “news platforms” will be used to describe any online media platform that involves delivering news content. e.g. Google News, Apple News

News platforms today have widely adopted algorithmic recommenders due to the explosion of digital content and competition amongst online media platforms (Thurman & Schifferes, 2012). Content recommenders utilise personalisation via machine learning to help order news content into user-friendly grouping, e.g., time, dates, geography and preference-based information. The mechanisms of these recommenders vary based on the different features extracted. Algorithmic personalisation, for example, utilises mathematical functions to identify and classify users to predict the content that might match the user preference. However, the end goals are similar: to improve the accuracy of the predicted recommendations. From the users’ perspective, personalised news feeds help prevent overloading information and increase user satisfaction. This is attractive to online media companies because the increased user satisfaction that comes with personalisation also increases user engagement. This engagement can be monetised through ads based on time exposed to adverts, as well as reduce the risk of churn for subscription-based news platforms. Despite the convenience and efficiency, concerns on these technologies, such as *filter bubble* and black box problem have also emerged. (Liang, et al., 2014)

Filter bubble and Echo chamber

Filter bubble and *echo chamber* originally come from different fields of research but ended up as a cause and effect under the context of personalisation technology. The term *filter bubble* was first coined by (Pariser, 2012), defining it as a result of personalisation technology. He reported that audiences within the *filter bubble* could only receive a limited and tailored selection of information, without any clear explanation of the filtering progress. Some researchers believe that the *filter bubble* can put people in a more vulnerable position of being targeted by harassing and misleading content and could cause injustice and polarisation in political affairs and beliefs.

Likewise, the term *Echo chamber, believed to first appear in academic literature* (Hall Jamieson & Cappella, 2010) , was used as a metaphor for an enclosed media space which is likely to reinforce conservative ideology. Audiences within the *echo chamber* are repeatedly exposed to the curated content that forms an insulated community. In recent years, the term has been extensively adopted to describe the phenomenon in the digital space. Most commonly, it is used to describe the result of the personalisation that occurs within the online media context. Under this context, it is broadly defined as an insulated online media space where people consume algorithmic filtered media. Concerns are that people within the community could be overconfident with their perspectives due to the repeatedly reinforced opinions that the personalised content they consume actually reflects a complete and balanced perspective.

A much-debated question is whether the algorithmic results of personalisation are the direct cause of societal issues, such as polarisation in political beliefs (Arguedas, et al., 2022). This concept has been recently challenged by researchers, suggesting more evidence and investigation is needed. Fletcher suggested that *filter bubbles* directly result from personalisation algorithms, whereas an *echo chamber is a hypothetical phenomenon requiring more evidence to understand its fundamental cause.* Other researchers believe that the problem that comes from a *filter bubble* should be made clear between self-selected personalisation and pre-selected personalisation. (Zuiderveen Borgesius, et al., 2016). They stated that humans have a natural tendency to seek opinions that are easy to digest because there are limitations for humans to absorbing a massive amount of information. In reality, humans will inevitably prefer supportive information over challenging ones (Dahlgren, 2021). Therefore, the arguments for demolishing news personalisation should not be made under the flawed premise. That is, to compare the online environment with the over-idealised world where people do not make any decisions (Fletcher, 2020). These scholars believe that users remain autonomous in selecting news sources and platforms despite automation within search engines and online news environments. Thus, research on this matter should not exclude the complex nature of societal issues.

Algorithmic Transparency

Despite the divergent opinions on the impact of *filter bubble*, both sides agree that more information about the technology should be introduced to the public. That means more transparency in the algorithmic process. In *Understanding Modern Transparency* (Meijer, 2009), the author defines transparency as "lifting the veil of secrecy". They addressed that modern transparency is mediated mainly through technology and computers when information can be mass produced and instantly distributed to individuals at a low cost. More factors are added between the source of content and the end-users. Technologies such as auto-generated content and algorithmic recommenders have convoluted the content creation process. Several existing recommenders are applying unexplainable machine learnings for the sake of accuracy, often referred to as "black box" models (Rudin & Radin, 2019). The black box problem has aggravated the path to reaching algorithmic transparency compared with the past. From the post-modern perspective, for average users, the most recent realisation is thus whether "computer-mediated transparency" fully reflects reality. The aggravated process of understanding the underlying mechanism of algorithms thus requires more information on, for instance, "How does the algorithm decide which content to provide to the users?", and "Why is certain information generated." That being said, methods for bridging the gap to answer these questions are needed.

The opaque process within the computation also led to difficulties for regulators. Incomplete laws had provided opportunities for misinformation and microtargeting on users to consume contents that may be harmful to the society out of unjustifiable business or political interest. (Zuiderveen Borgesius, et al., 2018) It's only recently that the EU has adopted GDPR as the first move towards explaining a fraction of personalisation algorithms. The protocol protected the individual's right to know how their data would be used. In the EU and the UK, the users now have some autonomy amongst the algorithms to control the exposure level of personal data. The GDPR policy has shown that regulations have helped justify the implementation of these automation algorithms and provided a more comprehensive range of choices for individuals (Gryz & Rojszczak, 2021).

Media and Information Literacy (MIL)

In a similar manner, approaches have been taken through education to empower individuals with critical thinking skills towards technology. Proposals on improving literacy towards personalisation technology have been vastly discussed. Various terms are coined, and besides conventional media literacy, educators have proposed "digital literacy" and "algorithm literacy" to specify the needs of these innovative fields (Anderson & Rainie, 2017) . So far, there is yet a standard term or definition that includes every aspect of the specific issue. Therefore, in this thesis, I adopted the term Media and Information Literacy (MIL) used by UNESCO. The term encompasses several aspects on media literacy along with clear definitions and structured frameworks. UNESCO suggests that MIL is fundamental and lifelong learning content. As communications are accelerated by the explosion of information and become more complex with the algorithmic applications, they urge institutions to adopt MIL as a set of competencies that every individual should acquire. Together these skills will build up the individual's resilience toward harm in new digital landscapes. (unesco, 2022)

Public Awareness and Needs

The need for MIL is not just a statement amongst governments and academia. According to a report, online users can be overconfident in their MIL ability, and because of that, users are less likely to check the source of the information they consume. Despite this overconfidence, the result is positive when asked whether it is essential to know more about the technology. A statistical result in the report by Ipsos MORI and Google indicated that approximately 66% of users believed that internet and technology companies should provide training to improve user's critical thinking." (Department for Digital, Culture, Media & Sport, 2021) Another piece of research focused on algorithmic transparency in the news media provided detailed information on how and what the users might learn about algorithms. They invited 50 academics and professionals in the news and media industry to an experiment to investigate their interest in knowing the mechanism and algorithms within mass media. These participants were particularly interested in knowing how their data was being used, and how their decision would affect the end product. The research suggests that interactivity with the end-user is a pivotal mechanism to help them understand the underlying algorithms. (Diakopoulos & Koliska, 2016) Other events show that people have gradually realised how algorithm change can impact what they receive. For example, in 2016, Twitter announced its change in adjusting the timeline from chronological to a preference-based recommendation. Twitter's decision has caused outrage from the users that invoked hashtags such as *#RIPTwitter* and *#F*ckTheAlgorithm* around this specific issue (Benjamin, 2022).

Intervention from the Academia

Researchers and institutions have proposed different approaches to intervene with the *filter bubble*. Many methods correspond to the MIL framework, as the approaches are often about informing users on media production knowledge or bringing up awareness of online events. The more people understand the process of news media production, the more likely they can identify disinformation and seek different media sources as alternatives. MIT Civic Media Lab proposed a media feed called *Gobo*, an ideal example of the intersection between MIL and explainable machine learning. The model they created aimed to increase user autonomy on the filtering algorithms by enabling the users to manage their filtered content. They suggested this model could improve transparency towards the algorithm and add perspectives outside the user's network, which could potentially help users to break the *filter bubble* (Bhargava, et al., 2019).

Aims of the Research

Most media studies exploring the solution of *filter bubble* have primarily focused on generating diverse and balanced content for the users. Others have tried to create more customisable filtering options for the users. Besides, in the machine learning field, research on personalise recommenders mainly focuses on improving the result of the algorithm but rarely on providing explanations to the users. With the exception of MIT Civic Media Lab, existing studies tend to give little attention to visualisation and interactivity as a medium for improving MIL.

To determine which part of the personalisation algorithm should carry out in the prototype, I referenced the preliminary work undertaken by (Diakopoulos & Koliska, 2016). Throughout their interview and discussion with the participants, they made a list of factors within the algorithmic systems that they believed were important information to be disclosed and explained. Specifically, for the algorithmic model section, they listed out factors such as "Input variables and features", "Sampling method", and "Ongoing human influence and updates". Additionally, interactivity is considered key for understanding the underlying information.

Therefore, in this thesis, I explore a creative approach to inform the users of the factors through visualisation. Thus, raising awareness for MIL and bridging the gap between the algorithm and humans. This thesis proposes an interdisciplinary method to make algorithms more approachable for average users. My aims are as followed:

1. To build a user-based model that will provide insight on user's relationship with the recommender. The model will be done through a process of collecting, storing and training via collaborative filtering.

2. To build a visualisation prototype built via p5.js that would enable users to interact with their positions within the personalisation system.

3. To assess these solutions' usefulness by surveying users familiar and unfamiliar with data science.

Methodology

Stage 1: User recruitment | Data Collection

The dataset used in this project is Microsoft News Dataset which is primarily used for news recommendation research (Wu, et al., 2020). The dataset contains over one million users with complete metadata on click history and impression logs. It also contains rich textual data with over 160K items of English news with labelled titles, categories and abstracts. For better training efficiency, I adopted a smaller version of the dataset which is a sample of 50K users with their digital footprints.

For the purpose of emphasising the idea of selecting behaviour, I opt for utilising the *Impression log* from the dataset. The *Impression log* is a series of data capturing the content shown to the users. In addition to this feature, the dataset also captured the click activity of the impressions. The click activity is in the binary format of news ID following a one or zero. One means the users have clicked on the news, and zero means the user had seen the news impression but did not click on it (Möbius, n.d.)

Due to the privacy policy for data ethics, the user data in Microsoft News Dataset is entirely anonymous, making the original users untraceable. Therefore, for qualitative evaluation later in the process, the model requires a new set of users to provide clicking data. The data collection was done separately through Google survey prior to the development of the prototype. There are two main benefits of using this method. First, it is more evident to the users. Second, it creates awareness for new users that their data is being collected.
I developed a survey to simulate a user's interaction with a news media website. I manually selected the news items to reduce the complexity of another layer of computer intervention. I narrowed the set down to four categories: News, Sports, Health, and Lifestyle.

Further, to emphasise the clustering result, the users were given a sample size of ten out of each news category randomly generated through a Numpy library function. The users would then select the news titles based on their preferences. These selections are then manually turned into click events. In alignment with the original dataset, the selected news titles are labelled as one and the rest as zeros. The advantage of this method is that it not only controls the impressions shown to the participants but also avoids a lengthy survey, which can help increase the participant's willingness to participate.

An important aspect of the data collection process is that the it is a one-class feautre and are collected manually. The purpose is to reduce the complexities that might come with automation. The reason is explained in contrast to other studies. In the research paper on Deep Structured Semantic Model (DSSM) (Huang, et al., 2013) , one of the recommended models by the Microsoft News Dataset is used for the content query technique in the research. It applies a deep neural network (DNN) for extracting semantic features from raw input text. This extraction is used for selecting candidate news articles for the users. However, automatic feature extraction as such would add more unexplainable factors to the model, such as why specific text features are selected but others are not. Additionally, recommenders that considered multiple features at a time, such as taking textual features into account, might distract the focus on user behaviour alone in the thesis.

Stage 2: Building NCF recommender model

The term "model" will be used to referred to the modified model under the NCF framework.

The thesis's interdisciplinary nature has impacted the selection of elements to build the model. As previously mentioned, the aim is to demonstrate a visualised prototype to make the personalisation algorithm phenomenon — "clustering" comprehensible to the average user. Therefore, the building blocks of a machine learning model, including the dataset, structure and visualisation, are selected primarily based on the end-user understandability. When searching for the appropriate recommender framework, the priority was reducing the complexity and noise that could potentially complicate the model.

As previously shown, the *filter bubble* has been criticised for high personalisation within machine learning. I searched for quantitative evidence within recommendation algorithms to understand the link between personalisation and such qualitative outcomes. Research has shown that collaborative filtering-based algorithms can cause less diverse recommendations over time (Nguyen, et al., 2014) . Collaborative filtering is one of the most successful mechanisms for recommendation systems. The method was inspired by humans' word-of-mouth communication, which seeks to mirror natural human behaviour. In real life, humans often make recommendations based on criteria for their subjective evaluation. These criteria can include a quality of an item, features of a service or others' comments on an event. In a similar manner, on collaborative filtering, the algorithm relies on users' ratings, information and content as representative features for training the machine. The machine then makes predictions based on what it calculated best for the users.

Note that my decision to apply collaborative filtering has not considered the long-term result it could cause on the users, as it is not the primary concern in this thesis. Neither have I considered improving the algorithm to avoid the pitfalls associated with collaborative filtering. The main focus of this paper is to explain the filtering process of the algorithm to the users.

I adopted user-based collaborative filtering focusing on user-based training to emphasise the relationships between the algorithms and the users. The specific framework used is an adaptation of Neural Collaborative Filtering (NCF), which focuses entirely on the collaborative filtering setting. (He, et al., 2017) The framework was selected for simplicity as I wanted to reduce redundant automation that might cause errors. The code for the model is an adaptation from Keras (Banerjee, 2020), which is initially used for movie recommendations.

Stage 3: Embedding layers and User Embeddings

The embedding layers play a crucial role in the model. The model is constructed by four fully connected embedding layers, as shown in Image 1. There are two sets of vectors for the user and news inputs, each of which follows a bias layer. In Image 2, each input layer is fully connected to a 16-dimensional embedding layer. Their primary purpose is to compress the sparse input vectors into low-dimensional dense vectors representing the features in a numerical format. The values generated for embedding layers are called — trainable weights — which are learned during the training. These weights can be used to observe the relationship between each data point and produce recommendations. (Hwang, et al., 2012). In this adaptation, instead of adding more layers to the neural network, the algorithm calculates the dot product of the embedding vectors through *tf.tensordot* , a Python, Tensorflow function that essentially replaces matrix factorisation. (He, et al., 2017)

```
Model: "recommender_net"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
embedding (Embedding)        multiple                  1760

embedding_1 (Embedding)      multiple                  110

embedding_2 (Embedding)      multiple                  31232

embedding_3 (Embedding)      multiple                  1952

=================================================================
Total params: 35,054
Trainable params: 35,054
Non-trainable params: 0
```

Image 1 Model Summary

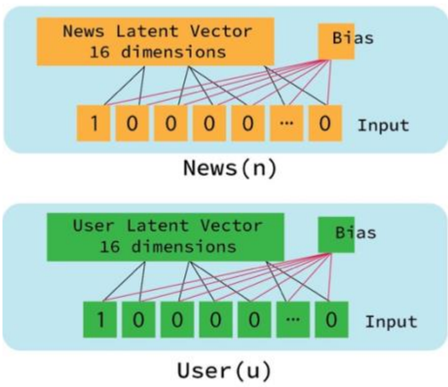


Image 2 The structure of embedding layers

Utilising one feature, however, could cause the one-class recommendation problem. The issue is known for suffering data sparsity and a lack of negative samples. Knowing that the binary data might affect the model's performance, the NCF framework considered binary properties as labels of how likely the news is to be relevant to the users. The researchers proposed a solution by applying log loss for the training. It has been suggested that the method can improve the performance of recommendations. Therefore, in this model, the training loss function is applied with *BianaryCrossentropy*, which according to the documentation on Keras, is explicitly used for binary classification applications. (Chen, et al., 2019). In other words, the model can interpret as a classification task for determining the similarity between the users based on their interaction with the news items.

The weights in the embedding layers help observe relationships between the data points. In order to demonstrate the relationship between the clustering users later in the process, the user embedding, I extracted the embedded weights in the user embedding layer to get numerical representations from the latent space. Image 3 is an excerpt of the weights from the user embedding. These floating numbers are trainable weights, meaning they could change every time the data is trained. However, this issue does not affect the primary purpose of communicating the user position in the 3D space. Hence, it is not the primary concern of the thesis.

```
[ 0.09879911  0.10543843  0.0859413   0.16353033  0.23674382  0.23887478
 -0.06299917  0.01596897  0.14329918  0.22484143 -0.19895294  0.16889562
 -0.01382881 -0.06523138 -0.01105244 -0.13172893]
```

Image 3 Embeddings in the 16-dimensional space

The prototype

Stage 1: Wireframe | Prototype

I initially looked for existing tools to enable me to visualise the embeddings. I explored using an open-sourced visualisation tool from Google — *Embedding Projector*. The tool provides a dynamic and intuitive dashboard for exploring the discrete structure of embeddings. It allows the user to visually observe the multi-dimensional space and offers standard dimensional reduction techniques such as Principal Component Analysis (PCA) and t-Distributed Stochastic Neighbour Embedding (t-SNE). Both are used for reducing high-dimensional data into low-dimensional space (Smilkov & , 2016) . *Embedding Projector* provides flexibility for creating non-default models. However, it lacks accessibility for average users. The customisable functions include dataset upload, customisable parameters for dimensional reduction and a draggable scale for the number of nearest neighbours. All of which would require a fundamental understanding of data science or machine learning to be able to navigate the tool. Thus, I chose not to use this for my primary data visualisation. Instead, I used this as a supplementary tool to explain the clusters in video format.

In search for an alternative, I used p5.js, a JavaScript language for creative coding, to create an accessible and interactive prototype. Several considerations have been put into the design process, such as information quality, interactive navigation, and accessibility to average users.

The idea of the visual demonstration lies between the *Embedding Projector* and pure textual description. The design of the prototype aims to include the users to be part of the decision process including inputting the data and reviewing the output of the training process. Each user was given a unique user id that was assigned to their data. Once they type their ids into the prototype, the data point following with numerical ids would light up in red in the space. Besides seeing their positions within the 3D space, the users can use their mouse cursor or fingers, depending on the device they were using, to zoom in, drag and turn the space in various angles to observe the space. Image 4 is screenshot of the prototype demonstrated from different angles.

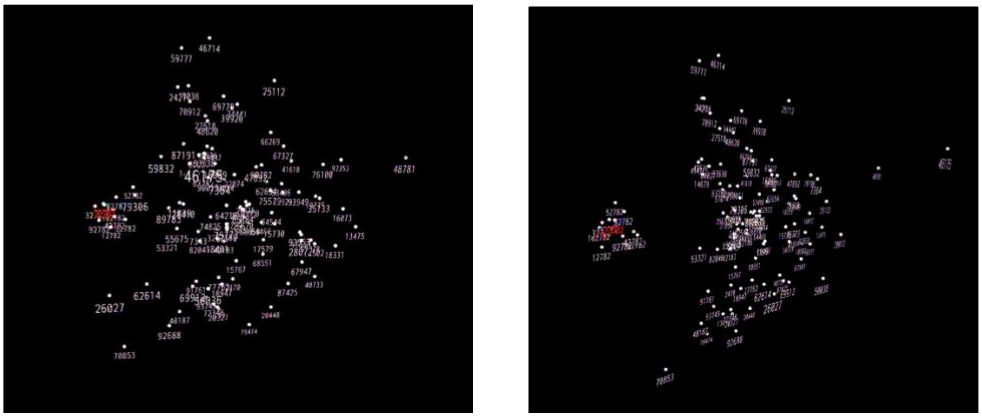


Image 4 Screenshot of the prototype

During the process of visualisation, I discovered that p5.js works at its optimal with around 150 data points. Any additional data would cause glitches and drastically increase the processing time. Therefore, the original dataset size was reduced to a sample of a hundred, that would leave some space for up to 50 more participants' data to be added in. This limitation of p5.js, however, allows a clearer vision on the data points. It is beneficial for user's visual interpretation, as it reduces the information in the space. A user would still be able to realise the clustering effect and their relationship with other users, despite the number of data point has been massively reduced. Another limitation of the 3D prototype is the lack of additional explanation to be informative. Within the 3D interactive space alone, there is neither descriptions nor instruction about the prototype. Hence, textual information is added for users who has no fundamental knowledge of machine learning.

The information cards aim to prepare the users with more knowledge on the topic before they start interacting with the p5.js prototype. The textual information is written in nine short passages, given that the longer text length might impact the users' interest. Additionally, for accessibility through different digital devices, the short passages are designed in the format of information cards. The cards are framed similarly to mobile device screen sizes to provide a better experience on smaller screens. Following the completion of textual content, to combine the Embedding Projector demo, the information cards and the prototype together, a wireframe for the webpage is presented in Image 5.

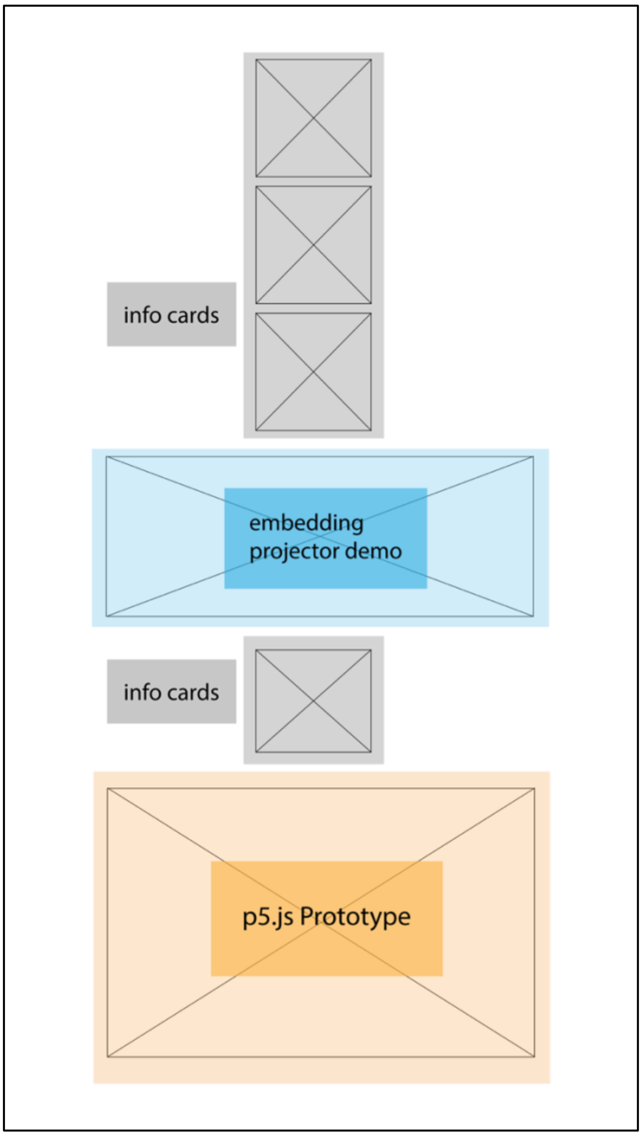


Image 5 Webpage wireframe

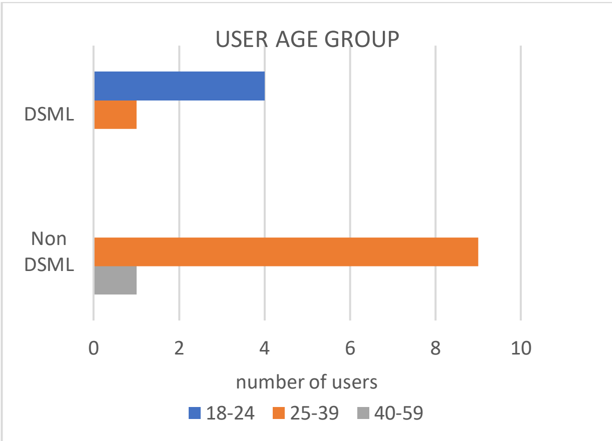
The descriptions begin with introducing the definition of personalisation in machine learning and the collaborative filtering approach used in automated predictions. I tried to avoid technical terms but to use plain language to describe the functions. The purpose is to reduce confusion from the users. Secondly, I introduced how personalisation technology is applied on news platforms. Later, I further emphasised the importance of MIL with supporting evidence from the UK government and UNESCO. These information cards provided a general understanding of the academic fields that are covered in the thesis.

The forth information card is about *echo chamber*. The term *echo chamber* was used as it is a term likely familiar to users, with the controversial part of term not addressed. The users have also learned about my proposal on building a prototype to communicate the concept through digital interaction. Just before the users watch the video clip of the Embedding Projector, they were informed on how to interpret the clusters and data points in the 3D space. Lastly, instructions were given in the last information card, to ensure that the users understand how they should interact with the prototype.

Stage 2: Survey design and result

The survey aims to evaluate whether the visualization model is helpful to increase users' understanding of personalisation in modern media production. To investigate the question, the participants were divided into two groups based on their knowledge level on data science and machine learning. Two groups of subjects were recruited, namely DSML participants and non-DSML participants.

- Participants who have studied or worked related to data science and machine learning (DSML participants)
- Participants who have NOT studied or worked related to data science and machine learning (Non-DSML participants)



In order to identify how well the participants understand data science and machine learning, they were asked to answer a series of questions on how they are familiar with some jargons, such as machine learning, personalisation, and media literacy. To assess the various aspects of users' experience, the questions are divided in four sections which are the following:

- *Informative*: How well the prototype has addressed the topic.
- *Usefulness and Importance*: Whether the participants consider the topic important and if the prototype is useful to them.
- *Change of Behaviour*: Whether the participants have changed their attitude towards the topic or would like to seek alternative source.

Each category contains several statements of opinions and experiences. The survey utilises the 7-point Likert scale to measure the participants attitude towards the statement. As Likert-type data are ordinal data, common statistic approaches such as mean, median, standard deviation, and t-test are not applicable because these measurements rely on the premise of normal distribution. Therefore, in the following assessments, I applied non-parametric analysis — chi-square test as an alternative (Mircioiu & Atkinson, 2017). Additionally, considering the sample size is small. I also applied Cramer's V test, as the supplementary evidence to show the power of the association between the two categorical fields (DSML and non-DSML). Apart from investigating the relationship between the two subject groups, I also evaluate the participants as a whole by interpreting the ratings through proportion.

The null hypothesis in this evaluation is that there is no relationship between the agreement level of the statements and the knowledge level in the professional field. If the null hypothesis is true, it means that the webpage has fulfilled its purpose according to the statement. By applying chi-square test, the result would show whether the null hypothesis should be rejected considering the different nature of the DSML participants and Non-DSML participants. Significance levels were set at 5% and the p values were generated for all variables.

During the statistical analysis, a Type II error occurred in the Topic Familiarity section. As previously mentioned, the controlling factor of two different groups should result in being significantly different towards the responses on their familiarity on the machine learning subject. However, due to the small sample size, it affects the analytical performance under large degree of freedom. Therefore, when conducting analysis based on the 7-point Linkert scale, chi-test failed to perform the correct significance level. For the purpose of reducing the degrees of freedom, I grouped up the variables in three categories, which were turned into Table 1:

Level of agreement	Ratings
Low	1, 2, 3
Medium	4
High	5, 6, 7

Table 1

Having the Type II error fixed, the next section will be demonstrating and interpreting the results from the user feedback. First, in the *Topic Familiarity* section, three questions were asked to ascertain the composition of the user groups. The questions are as followed:

*TF1: how well do you understand **machine learning**?*
*TF2: how well do you understand "**personalisation**" in a machine learning context?*
*TF3: how well do you understand the term "**media literacy**"?*

Topic Familiarity				
user	rate	TF1	TF2	TF3
DSML	low	0%	20%	20%
	medium	40%	0%	20%
	high	60%	80%	60%
non-DSML	low	90%	80%	40%
	medium	10%	20%	30%
	high	0%	0%	30%
	D.F	2	2	2
	p-Value	0.01	0.01	0.42
	Significance	YES	YES	NO
	Cramer's V	0.63	0.58	0.20

Table 2 Topic Familiarity

In *Table 2*, according to the p-Value, **TF1** and **TF2** show that the null hypothesis is rejected. The results indicate that there are significant differences between the two groups which statistically justified the premise of two distinct groups of participants. This also proved that the Type II error has been solved. Meanwhile, for **TF3**, the result appears that whether the participants have studied data science or not has no statistical impact on their understanding on media literacy. The result helps normalise the setting of the evaluation, with both groups having similar understanding of MIL.

Moving onto the next section of the survey, the participants were asked to follow a link to the webpage, to read through the information cards, and to input their unique user ids into the designated space, until they see their id number highlighted in red. They were told to then come back to the survey and answer the rest of the questions.

The second section of question is on *Usefulness and Importan*ce of the prototype. The series of statements were designed to investigate if the participants consider the topic as important to them:

UI1: It is important for me to know how my data is being classified.
UI2: I find useful to see how my news choices relate to other people's.
UI3: Online news platforms should reveal users' data description on the webpage.

According to table 3, both sets of participants want to know more about the algorithms, with no statistical difference between the two. Based on the statistical proportion of ratings in each group, over 70% of the users believe that it is crucial to understand how their data is being classified by the algorithms. Moreover, 80% of the participants believes that it is useful to see their choices through the prototype in relation to others. Lastly, 87% of the participants believes that new platforms should reveal more discrete data to the users. Specifically, within the DSML group, all of the participants gave high responses to the statements. It can thus be concluded that the majority of the participants is acceptive towards the idea of learning more about personalisation.

Usefulness | Importance

user	rate	UI1	UI2	UI3
DSML	low	0%	0%	0%
	medium	0%	0%	0%
	high	100%	100%	100%
non-DSML	low	10%	10%	0%
	medium	30%	20%	20%
	high	60%	70%	80%
	D.F	2	2	2
	p-Value	0.192	0.392	0.562
	Significance	NO	NO	NO
	Cramer’s V	0.332	0.250	0.196
Overall	low	7%	7%	0%
	medium	20%	13%	13%
	high	73%	80%	87%

Table 3

The *Informative* section is on whether the prototype communicates the messages well to participants. The statements are as followed:

IN1: I feel much more informed about how personalisation is used in the news media.
IN2: I have a much greater understanding on how online news is produced.
IN3: I understand that news personalisation affects the news I see.
IN4: I understand that news personalisation can cause echo chamber.
IN5: It is surprising that my action on selecting the news affects the result of personalisation computation

According to table 4, there is a significant difference on statement IN1 while the rest remains similar. None of the participants in either group rated low for IN1, however, there's a 20% difference of the participants on whether they feel much informed on the topic. This result may be explained by the fact that the non-DSML cohort have less background knowledge on personalisation mechanism, causing them to feel they have gained more knowledge on the topic. For all other questions we can see that both DMSL & non DSML both feel more informed after viewing the webpage.

Informative						
user	rate	IN1	IN2	IN3	IN4	IN5
DSML	low	0%	0%	0%	20%	20%
	medium	40%	20%	40%	20%	40%
	high	60%	80%	60%	60%	40%
non-DSML	low	0%	0%	0%	0%	30%
	medium	20%	30%	10%	10%	0%
	high	80%	70%	90%	90%	70%
	D.F	2	2	2	2	2
	p-Value	0.010	0.92	0.34	0.27	0.1
	Significance	YES	NO	NO	NO	NO
	Cramer's V	0.835	0.08	0.26638	0.14	0.39
Overall	low	0%	0%	0%	7%	27%
	medium	27%	27%	20%	13%	13%
	high	73%	73%	80%	80%	60%

Table 4

The last section of the survey, *Change of Attitude,* elicited information on the participants' attitudes on whether the webpage has impacted their behaviour when interacting with news media. For example, becoming more critical of personalised news content or attempt to diversify their sources. The statements in the survey are as followed:

CH1: I want to understand more about personalisation used in news media.
CH2: I have changed my perspective on online news media.
CH3: I want to find ways to see different perspectives of the news.
CH4: I want to seek other sources/platforms for news media.
CH5: I want to know other people's selection of news articles.
CH6: I want to see more visualisations on how my data is being used by news platforms.

Change							
user	rate	CH1	CH2	CH3	CH4	CH5	CH6
DSML	low	20%	0%	0%	0%	0%	0%
	medium	0%	40%	20%	0%	0%	20%
	high	80%	60%	80%	100%	100%	80%
non-DSML	low	10%	20%	0%	20%	20%	10%
	medium	0%	70%	10%	10%	20%	0%
	high	90%	10%	90%	70%	60%	90%
	D.F	2	2	2	2	2	2
	p-Value	0.866	0.011	0.866	0.392	0.256	0.281
	Significance	NO	YES	NO	NO	NO	NO
	Cramer's V	0.098	0.534	0.053	0.171	0.244	0.291

Table 5

Amongst all of the statements, in table 5, only CH2 shows a significant difference. 60% of the DSML participants show that they have increased scepticism towards online news, whereas fewer people (10%) from the non-DSML participants have agreed with the statement. These results suggest that audiences with deeper understanding of personalisation algorithms are more likely to be critical of news media.

Overall, the results suggest that data visualisation combined with information cards can show within my webpage was successful in communicating the concept of personalisation within news media to the participants. With only a few statements (23%) rejecting the null hypothesis we can see that a prior knowledge of personalisation algorithms is not necessary to become literate in MIL.

The final section of the survey is comprised of textual user feedbacks on the webpage and their opinions on personalisation technology. From the non-DSML group, participants commented that they are not sure if they would prefer to receive even more information about the source of the news, because it is already overwhelming even just consuming news online nowadays. Similarly, there are participants that don't mind consuming filtered content, but would prefer more authentication on the news content. "I think fake news on the internet are inevitable, so the big websites like Facebook or Google should have a system to detect suspicious contents and kindly remind the users.", according to one of the responses. Whereas some participants find personalisation gives them a sense of being unconsciously manipulated. Participants also proposed if there are existing tools that can avoid being categorize by the algorithms.

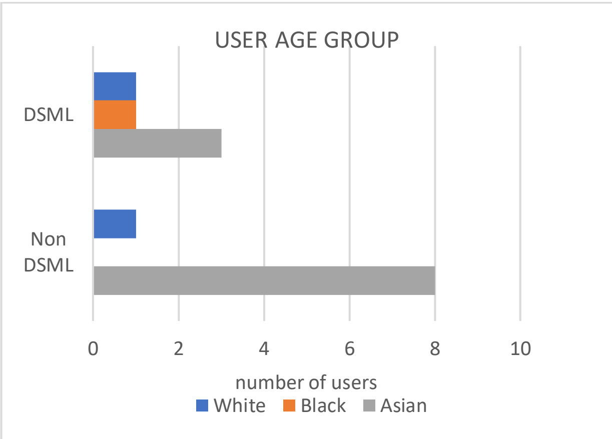
On collecting feedback about further development of this prototype, the participants suggest that it would be nice to reveal more information about the other users, such as their countries and regions. "Some discussion of fake or misrepresenting news and how that affects users would also be useful.", one of the participants respond to the survey. Especially for DSML participants, they tend to ask for more information on what could be further added onto the prototype. Such as more description on the news' content as well, showing the result of what genre or keyword have been clustered together. In this section, DSML participants ask for more information and features to the prototype, on the other hand, non-DSML participants prefer less information but more justifiable content.

Limitation | Future Work | Suggestion

One of the limitations of this study is that several parts of the personalization model remains unclear due to the "black box" problem. While the concept of collaborative filtering is designed to be simpler to comprehend, it is still considered as a "black box". The reason is that it largely relies on sparse and incomplete data, (Herlocker, et al., 2000) therefore, it is still unexplainable on how the user embedding parameters are generated. Whilst, this webpage offers valuable insights on the possibility to communicate and raise awareness on MIL via demonstrating users in a visualised embedding space, this study did not explain the full mechanism of the algorithm.

Secondly, the sample size for evaluating user feedback is smaller than the recommended size for statistical analysis. I attribute the problem to the long period between collecting the user's data and them completing the follow up user survey. 25% of the participants did not return to participate in user experience feedback after they participated in the first one. Chi-square test is a sample size sensitive method; therefore, it is possible to suggest that the result might be different with a larger sample size. I suggest two solutions for this issue. Firstly, include more users to the survey, that even if a proportion of them did not participate for the user feedback, the project would still have a reasonable sample size for analysis. Second, a one-site input and output interface would remove the need for a delay. Users could select their preferred news content and see the outcome of the 3D prototype instantly, which would fill up the gap between different stages of the project.

Lastly, the selection of the user group should be more inclusive. The chart below is the ethnicity statistic from the survey. The survey candidates were predominantly fellow UAL students and the diversity of the survey hence reflected the diversity of the group reached. This also caused the lack of the vulnerable users within the user group. According to the DCMS (Department for Digital, Culture, Media & Sport, 2021), those users are often the most impacted group when it comes to misinformation and online abuse. This can be solved by collaborating with vulnerable users working group or education institutes.



For future work, I plan to complete the prototype where the data collection can be implemented within the same page / interface of the interactive 3D prototype. This could be helpful to make sense of the connection between the data collection and the visual outcome. Further research on explainable artificial intelligence would be great to understand more on technical aspects of approaching algorithmic transparency, continuing on making algorithms more accessible to average users.

As regards of suggestions for other implementation of the method proposed in the thesis, approaches for dimensional reduction should be carefully considered for its meaning. t-SNE was once considered as one of the approaches in the thesis. However, t-SNE is a non-deterministic algorithm with changeable parameters such as perplexity and learning rate. t-SNE produces random results every time it is initialised. In comparison, PCA works as deterministic algorithm, which the result is the same given the same input. Therefore, I suggest that before a better explanation on t-SNE is introduced, a study similar to this one should be carried out using PCA as the dimensional reduction method. (VIÉGAS, et al., 2016)

Conclusion

This study set out to explore whether visualizing the algorithmic structure can increase the audiences' awareness and understanding of personalisation in news media. The interdisciplinary research method used has revealed new possibilities on the intersection of machine learning, media and information literacy (MIL) and visual communication. The MIL framework emphasizes the essentiality of developing tools and methods to help people becoming resistant to the negative impact of personalised media content. In this research, I focused on using the example of personalised news, which is a main tool of how people consume news content nowadays (Fletcher, 2020). A lot of the users are unaware of impact from the technology, thus putting the themselves in vulnerable positions, unable to critically identify the intention or source of the news. The machine learning techniques helped accomplish the experiment. I build up a modified NCF recommendation model, that generates user embedding parameters based on the existing and newly collected user data. This allowed me to further apply PCA for reducing the high-dimensional space to a 3-dimensional space that humans can visually comprehend. Lastly, with p5.js, I build up an interactive prototype along with textual information to demonstrate the data points to the audiences. According to the user feedback, DSML and non-DSML participants both showed positive attitudes towards learning more about personalisation of news production, the ability to make sense of algorithms as well as indicating that understanding of this algorithm generates awareness and inspires curiosity.

At last, while the prototype raised awareness around the potential negative impacts of personalised news media, the user survey suggests that receptive users can navigate the complexity of the change in modern news media production with a sufficient explanation. The prototype has improved communication and made the algorithm more accessible to general audiences and reduced the overwhelming anxiety of information that personalisation content would only do harm. A few decades ago, when the television replaced traditional newspapers and the radio and became the dominant mass media, there was criticism towards the invention and that even emerged an academic field on "television criticism", for developing appropriate language of its content and content production. (Lotz, 2008) Today, the younger generation rarely get deceived by TV commercials, and many of them understand there are different political inclination from different news channels. I believe the same with algorithmic recommenders. With education and more diverse presentation on the algorithmic personalisation, which I had proposed in this thesis as an example, would help us to establish a healthier and safer environment to interact with machine learning algorithms and personalization technologies.

Acknowledgements

I would like to thank the following people for helping with this research project:
My supervisor Dr. Louis McCallum, for his support and patient. He inspires me with his knowledge on the subject; For people who helped with the data collection and user feedback survey, the thesis would not have completed without your support; My best friend and partner Jack, who encourages me throughout the time when I doubt myself the most; And special thanks to my family, Patrick, Lisa and Wayne, thank you for supporting the decision on pursing the MSc degree.

Bibliography

Mircioiu, C. & Atkinson, J., 2017. A Comparison of Parametric and Non-Parametric Methods Applied to a Likert Scale. *Pharmacy,* 5(2), p. 26.

Anderson, J. & Rainie, L., 2017. *Code-Dependent: Pros and Cons of the Algorithm Age.* [Online]
Available at: https://www.pewresearch.org/internet/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age/ [Accessed 4 September 2022].

Arguedas, R. A., Robertson, C. T., Fletcher, R. & Nielsen, R. K., 2022. *Echo chambers, filter bubbles, and polarisation: a literature review.* [Online]
Available at: https://reutersinstitute.politics.ox.ac.uk/echo-chambers-filter-bubbles-and-polarisation-literature-review [Accessed 14 September 2022].

Banerjee, S., 2020. *Collaborative Filtering for Movie Recommendations.* [Online]
Available at: https://keras.io/examples/structured_data/collaborative_filtering_movielens/ [Accessed August 2022].

Benjamin, G., 2022. #FuckTheAlgorithm: algorithmic imaginaries and political resistance. *ACM Conference on Fairness, Accountability, and Transparency,* 20 June.pp. 46-57.

Bhargava, R. et al., 2019. Gobo: A System for Exploring User Control of Invisible Algorithms in Social Media. *CSCW '19: Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing,* pp. 151-155.

Chen, J., Lian, D. & Zheng, K., 2019. Improving One-Class Collaborative Filtering via Ranking-Based Implicit Regularizer. *Proceedings of the AAAI Conference on Artificial Intelligence,* 17 July, 33(1), pp. 37-44.

Cinelli, M. et al., 2021. The echo chamber effect on social media. *Proc Natl Acad Sci USA,* Volume 118.

Dahlgren, P. M., 2021. A critical review of filter bubbles and a comparison with selective exposure. *Nordicom Review,* 42(1), pp. 15-33.

Department for Digital, Culture, Media & Sport, 2021. *Online Media Literacy Strategy,* s.l.: s.n.

Diakopoulos, N. & Koliska, M., 2016. Algorithmic Transparency in the News Media. *Digital Journalism,* 27 July.
Fletcher, R., 2020. *The truth behind filter bubbles: Bursting some myths.* [Online]
Available at: https://reutersinstitute.politics.ox.ac.uk/news/truth-behind-filter-bubbles-bursting-some-myths
[Accessed 1 September 2020].

Gryz, J. & Rojszczak, M., 2021. Black box algorithms and the rights of individuals: no easy solution to the “explainability” problem. *Internet Policy Review*, 10(2).

Hall Jamieson, K. & Cappella, J. N., 2010. What Do We Mean by "Echo Chamber"?. In: *Echo Chamber.* New York: Oxford University Press, pp. 75-77.

Herlocker, J. L., Konstan, J. A. & Riedl, J., 2000. Explaining Collaborative Filtering Recommendations. *CSCW '00: Proceedings of the 2000 ACM conference on Computer supported cooperative work*, December.p. 241–250.

He, X. et al., 2017. Neural Collaborative Filtering. *WWW '17: Proceedings of the 26th International Conference on World Wide Web,* 3 April.pp. 173-182.

Huang, P.-S.et al., 2013. Learning Deep Structured Semantic Models for Web Search using Clickthrough Data. October.p. 2333–2338.

Hwang, Y., Han, B. & Ahn, H.-K., 2012. A fast nearest neighbor search algorithm by nonlinear embedding. *2012 IEEE Conference on Computer Vision and Pattern Recognitio,* 16-21 June.

Jolliffe, I. T. & Cadima, J., 2016. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A ,* 13 April.374(2065).

Liang, T.-P., Lai, H.-J. & Ku, Y.-C., 2014. Personalized Content Recommendation and User Satisfaction: Theoretical Synthesis and Empirical Findings. *Journal of Management Information Systems,* 9 December, 3(23), pp. 45-70.

Lotz, A. D., 2008. On "Television Criticism": The Pursuit of the Critical Examination of a Popular Art. *Popular Communication,* 6(1), pp. 20-36.

Möbius, n.d. *MIND: Microsoft News Recommendation Dataset.* [Online]
Available at: https://www.kaggle.com/datasets/arashnic/mind-news-dataset
[Accessed August 2022].

Meijer, A., 2009. Understanding Modern Transparency. *International Review of Administrative Sciences,* 75(2).

Nguyen, T. T. et al., 2014. *Exploring the Filter Bubble: The Effect of Using Recommender Systems on Content Diversity.* Seoul: Association for Computing Machinery.

Pardos, Z. A. & Jiang, W., 2019. *Combating the Filter Bubble: Designing for Serendipity in a University Course Recommendation System.* [Online]
Available at: https://doi.org/10.48550/arXiv.1907.01591
[Accessed September 2022].

Pariser, E., 2012. *The Filter Bubble: What The Internet is Hiding From You.* s.l.:Penguin.
Poggenphol, S. H., 2018. *Design Theory to go.* Colorado: Ligature Press.

Ricci, F., Rokach, L., Shapira, B. & Kantor, P. B., n.d. *Recommender Systems Handbook.* New York: Springer Science+Business Media.

Rudin, C. & Radin, J., 2019. *Why Are We Using Black Box Models in AI When We Don't Need To? A Lesson From an Explainable AI Competition.* [Online]

Available at: https://hdsr.mitpress.mit.edu/pub/f9kuryi8/release/8 [Accessed 22 September 2022].

Smilkov, D. & B. P. G., 2016. *Open sourcing the Embedding Projector: a tool for visualizing high dimensional data.* [Online]
Available at: https://ai.googleblog.com/2016/12/open-sourcing-embedding-projector-tool.html [Accessed 4 September 2022].

Thurman, N. & Schifferes, S., 2012. THE FUTURE OF PERSONALIZATION AT NEWS WEBSITES. *Journalism Studies,* 27 March, 13(5-6), pp. 775-790.

unesco, 2022. *About Media and Information Literacy.* [Online]
Available at: https://www.unesco.org/en/communication-information/media-information-literacy/about#:~:text=Media%20and%20information%20literacy%20is,information%2C%20digital%20and%20communication%20landscapes.
[Accessed September 2022].

VIÉGAS, F., WATTENBERG, M. & JOHNSON, I., 2016. *How to Use t-SNE Effectively.* [Online] Available at: https://distill.pub/2016/misread-tsne/#citation
[Accessed 19 September 2022].

Wu, F. et al., 2020. MIND: A Large-scale Dataset for News Recommendation. July, Volume Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, p. 3597–3606.

Zuiderveen Borgesius, F. J. et al., 2018. Online Political Microtargeting: Promises and Threats for Democracy. *Utrecht Law Review,* 14(1), pp. 82-96.

Zuiderveen Borgesius, F. J. et al., 2016. Should we worry about filter bubbles?. *INTERNET POLICY REVIEW Journal on internet regulation,* 31 March .5(1).