

How to burst your bubble: Broadening your social media horizons

February 4 2021



Credit: [NeONBRAND](#) on [Unsplash](#)

Evangelos Papalexakis is an assistant professor of computer science and engineering at UC Riverside's Marlan and Rosemary Bourns College of Engineering. His research spans data science, signal processing, machine learning, and artificial intelligence. One of his ongoing projects aims to develop an automated fake news detection mechanism for social media.

Most people know by now that what they see on [social media](#) sites like Facebook has something to do with mysterious algorithms. Can you explain what algorithms are, in general?

You can view an algorithm as a set of instructions that a computer has to follow to solve a problem, much like a recipe where the input is the ingredients and the output is food. The algorithm has inputs, which could be data, and outputs, which could be the solution to a problem.

Another term we see a lot is machine learning. Can you explain what that is?

Tom Mitchell in his classic textbook defines [machine learning](#) as the study of algorithms that improve their performance on a particular task through experience. Experience usually refers to data in that case.

Frequently, we refer to a machine learning model as the product of a machine learning "training" algorithm whose job is to learn how to solve the particular task assigned to it given the data, and then distills that knowledge in a model, which can be as simple as a set of IF-THEN-ELSE rules, to something as complicated as a neural network.

After training, we deploy the machine learning model and it then follows another algorithm, typically called an "inference" or "prediction" or "recommendation" algorithm which, using the existing model and given a particular user, outputs the content, sometimes in a ranked list, that the user is most likely to engage with.

How do social media companies use machine learning models to filter what shows up in our feeds?

In this particular case, the task is to figure out what to show to a user. The experience is all the user's interaction/engagement/content creation in the platform, and the performance can be measured by whether the user successfully enjoyed and/or engaged with the recommendation—the item shown in the feed—or not.

Netflix pioneered this thing by starting a competition that had this exact task in mind and entailed a monetary prize. In the solution that won the competition, and basically any recommendation machine learning algorithm, everything boils down to computing a "representation" of a user and a "representation" of the content, and then figuring out which type of content, such as a movie, a certain user is most likely going to enjoy.

In simple terms, imagine user representation as an Excel spreadsheet whose rows are users and the columns are different movie genres and each cell telling us how much this user "prefers" this particular genre. If we use a similar representation for the movies, then we can basically see which user has a high match with which movie, in that "genre" representation. The key now is to identify those "genres" from the rich amount of data in the platform. The genres given by the movie studios don't necessarily reflect the context in which people watch them, but emerge from the patterns of users interacting and consuming content. Similarly, social media platforms use data created and shared by a user and all the kinds of interactions that user has with other content creators or with content to assign their own "genres."

Data is the gamechanger. All research in machine learning is open and shared publicly at a very rapid pace, both from academia and industry. What makes a difference is the data used to "train" the machine learning models. Anyone in the world is able to experiment and tinker with the most cutting-edge models used, but without the same data to train it, and that data is really what makes the difference. In the case of social media, our online behaviors are the data.

How do Facebook's models drive people toward groups, pages, and individuals who share their same interests, creating echo chambers?

In general, a "like" means positive engagement. Therefore this becomes a signal that is fed into the training of the model and used to update and refine the representation of the user, meaning, the set of preferences that the algorithm has learned for this particular user.

The machine learning models are aiming at determining what is the best next thing that a user would be most likely to engage with, for example, "like" or "give 5 star rating."

We cannot quantify exactly the effect of each kind of engagement, since it really depends on the specific model and how it was trained. For example, is "liking" the same as "sharing" a post? Is giving a two-star rating a stronger signal than watching the first 10-15 minutes of a movie and then quitting? But in general, it makes sense to expect that the more we engage with a specific type of content, such as comedy movies or pictures of dogs, the more we signal to the model that this is what we like.

Given that the model is trained with this as a primary objective, it will favor content that resembles content that the user has already engaged with, and this means that, in the vast ocean of content being shared in a platform, it will most likely rank other content lower.

What can people do to influence Facebook's machine learning models to show them more diverse content, whether it's dinner pictures from friends or national news?

It is unclear how that can be systematized, since there is no way of knowing exactly by how much each engagement influences the model and by how much it depends on the type of content, such as breaking news vs. pictures of pets. This points to the need for model transparency

which could provide a human readable summary of what the model thinks our preferences are and perhaps the ability to tweak them. Platforms sometimes do the latter by directly asking whether this content is relevant right now, which is something they obviously cannot do all the time, otherwise the user would be understandably annoyed. But good ways of being transparent are now a very important direction in the research community. You may have noticed Netflix, for example, sometimes says that, "Because you watched XYZ we recommend the following movies."

A major challenge in the above transparency direction is that the representations learned are usually with respect to "genres" that are not necessarily human-readable or at least not immediately insightful to someone by mere inspection.

Right. You mentioned earlier that in the example of movies, the genre given by the movie studio might be different from the genre the machine learning model assigns based on how we interact with the content.

A Netflix example could be a set of movies that have no apparent common thread between them other than the fact that people mostly watch them ironically, such as "The Room."

This is not an easy to define genre, but extremely helpful in understanding how real users enjoy content, perhaps differently from how their creators envisioned it. Going back to the spreadsheet analogy, the fact that those columns in the Excel sheet of the representation learned are not always intuitive, it is very challenging to provide a fully-understandable explanation or justification based on those columns. For example, because you have a high score in this category, Facebook shows you this post, but "this category" is not easy to describe in words.

It is more likely to be combinations of such categories, complicating the picture even more.

It sounds like one way to put a crack in a filter bubble is to diversify the way we engage with content to nudge the model in other directions?

I, personally, sometimes go out of my way to identify sources of content that perhaps my immediate online social circle would not share and engage with them so that I signal the model that this is part of the stuff I would like to see more of.

From the point of view of machine learning, we would have to define additional constraints on how to measure the algorithm's performance which would somehow encode diversity of content. This is a very challenging research problem especially given that it is hard to quantify that objective in a unique way.

There is a lot of fascinating research that talks about diversification of content recommendation and bursting the filter bubble, and it is, in fact, a very challenging research problem. An insightful Facebook AI blog post talks a bit about how this is done in Instagram's Explore function, where they are trying to discourage showing content from the same user very often in order to allow the algorithm to retrieve content from other accounts which can hopefully be a bit more diverse.

Do you have any recommendations for how people can learn to recognize and manage their own reactions and online behaviors so that they don't live in an echo chamber?

It is important to internalize that in all platforms that somehow rank their

content, what is being shown to us is only a small subset of what is out there. Focusing only on a subset is not necessarily a bad thing, since there is so much content competing for our attention that it would quickly deplete our attention and ability to learn or enjoy anything.

I view those ranking/recommendation systems as machine learning-based assistants that do their best to understand what I like, but that can also sometimes get tunnel vision and perhaps I can try to give them some more information a bit more deliberately. However, it is very important to understand that fact because if we conflate what we see in our feed with what we think is the totality of things being shared online, this can absolutely lead to filter bubbles.

In order to understand the impact that our online actions in a platform have on the content that we are served next, it is a fun experiment to try and, for example, like every picture of a dog but not cat for a few days and observe what happens to the posts you see after that. Subsequently, after a couple of days of observing any change, explicitly seek out accounts that share pictures of cats too and like cat and dog content an equal amount, and then observe how the recommendations you get change.

What we control in that system is the data we create and that is fed into the [machine learning model](#). So, if we are a bit more particular about how we create that data, this can also help nudge the [model](#) to increase the diversity of our preferences. How does this translate to practice? For instance, in breaking the filter bubble of news-related content that we see, it is a good idea to actively seek out reputable news outlets that span the ideological spectrum and follow them, thereby engaging with their content at-large, even though perhaps our immediate social circle may share things from a part of that spectrum.

Provided by University of California - Riverside

Citation: How to burst your bubble: Broadening your social media horizons (2021, February 4)
retrieved 25 April 2024 from

<https://techxplore.com/news/2021-02-broadening-social-media-horizons.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.