Janice Ainembabazi

Southern New Hampshire University

CS-370-12200-M01 Current/Emerging Trends in CS

Dr. Mohammad Habibi

November 23rd 2024

At the moment, most personalization systems rely on a class of machine-learning algorithms known as neural networks. Neural networks are a cornerstone of modern user-experience personalization. These systems are not quite as complex as the human brain, but they do have layers through which data passes and basic predictions are made. They take in data from an input layer, where the raw numbers that describe a user's behavior are collected (e.g., how long a user has been on a site, what he's clicked on, or the route he seems to be taking through a series of pages). Then the information goes to one or more hidden layers, where the real "learning" occurs. Here, the data is intensely analyzed, with various parts being assigned weights that describe their relative importance or salience. It comes out the other end, at the output layer, as a recommendation, a tag, a score, or some other sort of e-presentation. The strength of neural networks is their capacity for adaptation and improvement through a method called backpropagation. When a system makes a mistaken prediction, it learns from the error by adjusting the connections between the nodes of the system. Over time, this feedback loop helps the system become more accurate. Companies have started using the technology for delivering user experiences that are not just better, but also more personalized. Yet this comes with some troublesome challenges, especially when the technology is applied in ways that involve collecting and processing large amounts of data.

Neural networks are widely employed in personalization, and for an excellent reason. They work astoundingly well in analyzing huge amounts of user data. When it comes to user engagement, the more a system knows about a user, the better it can serve that user. And neural networks are really good at this, that they're very difficult to beat. Earning a solid engagement with a user is an extremely reliable way of increasing the chance that a user will click on an ad, and this is a reliable way of making money. There are, nevertheless, ethical concerns related to personalization. Neural networks frequently function as "black boxes," with no clear reasoning as to why they make the decisions they do. Who really knows, for instance, why the network decided to recommend a certain item to the user or serve them a particular ad? Because we pretty much have to take the network's word for it, this opacity, that is, the nature of being not easily seen through or understood, can be an ingredient in the slow-baking trust casserole. Which poses questions like: You want the network to be fair, right?

The European law known as the General Data Protection Regulation (GDPR) aims to safeguard the data and privacy of individuals. It directly impacts our use of neural networks for personalization, obliging us to conform to several of its key principles. The first, of course, is that we must have a lawful basis for the processing of personal data. The most relevant for us at the moment is "consent", before we can use your data to inform our models of what might be a good personalized experience for you, we must first obtain your permission. A basic requirement is that companies must be clear and direct about what they do with user data, purpose limitation. They must say how they process it and for what purposes promoting transparency. Another key part of the regulations is that companies can only use the data for the specified purposes. If they have some new idea, they can't go ahead and use the data for it unless they first get permission from the users for accountability. Data collected must be accurate, and if it's not, it must be made accurate in a timely manner. Then there are the rules about keeping confidentiality, requiring that strong measures be taken to prevent unauthorized access to the data. If a company does all this and acts in a trustworthy way, enough so that if a user happened to be around when the company was processing the user's data, the user wouldn't be shocked or appalled, then the company is probably in compliance with the rules. These principles directly affect our company’s usage of neural networks. Take the collection and storage of user data. Right now, we collect and store lots of user data and we store it for a long time. But what we’re doing is almost the opposite of what the GDPR says to do. And when it comes to using that data, the GDPR requires user consent for us to use data in ways that are even somewhat similar to our original purposes. It’s a good rule. But if we tell users what we’re doing, I bet many (if not most) would be very uncomfortable with the whole thing.

Even with these worries, it is not an option to abandon entirely the collection of data for our business model. To work well, our systems of personalization require user data, without which we would lose our competitive edge. So, instead of going data-free or even "almost data-free," we are working on acquiring data in ways that maintain the required level of functional and valuable "neural networks" that power our systems of personalization. In short, we are looking for acceptable ways and means to acquire data. We have several routes to explore like these two; One is using privacy-preserving methods that the Stanford researchers work with, most notably differential privacy. A further significant advance that can be made is to adopt explainable AI, which aims to make neural networks clearer and more user-friendly. If we want to build trust among users and comply with the "transparency" part of the GDPR, then we need to get better at saying why we made a given recommendation. And these days, with privacy put under a microscope, it's just not enough to say we've got all the right "server-side" protections. We also need to be able to demonstrate, through regular audits of our data practices, that mainframe-committed orders don't keep our datasets for unacceptably long periods of time.

**REFERENCE**

Bhatt, D. (2022, February 10). *Privacy-preserving in Machine Learning (PPML)*. Analytics Vidhya. Retrieved November 23, 2024, from <https://www.analyticsvidhya.com/blog/2022/02/privacy-preserving-in-machine-learning-ppml/>

DeepAI. (2019, May 17). *Hidden layer*. DeepAI. Retrieved November 23, 2024, from <https://deepai.org/machine-learning-glossary-and-terms/hidden-layer-machine-learning#:~:text=In%20neural%20networks%2C%20a%20hidden,inputs%20entered%20into%20the%20network>

European Union. (2016). General Data Protection Regulation (GDPR). Available at: <https://gdpr-info.eu/>

*What is an input layer? - definition from Techopedia*. Techopedia.com. (n.d.). Retrieved November 23, 2024 from <https://www.techopedia.com/definition/33262/input-layer-neural-networks>