Project One

Name: Matthew Rix

Class: CS-370-T3355

Date: 1/27/2023

Here at Social Media Company (SMC) we pride ourselves on delivering a one-of-a-kind user experience personalized to our users to connect and share their life with those they love and their community wherever that may be. In order to provide such an experience to our clients we here at SMC utilize the cutting edge of digital technology. One of these technologies is called neural networks which allow us to create personalized models for each individual user, but you may be asking yourself what is a neural network?

Just as neurons transmit and process data in animal and human minds (Byrne 2016); neural networks based on biological neurology utilize artificial neurons in order to transmit and process data in the same way. These neural networks are composed of multiple layers known as the input, hidden, and output layers. The input layer brings the initial data into the system for further processing by subsequent layers of artificial neurons.

The input layer is the very beginning of the workflow for the artificial neural network (Techopedia 2023). The next layers are the hidden layers in which the algorithm applies its function to the inputted data from the input layer (DeepAI 2019). It should be noted that the moniker ‘hidden layer’ is simply a term which illustrates that it is not a portion of the neural network that the user interacts with. Finally, the output layer produces the final prediction results.

The result of the neural network is to create the personalized experience our users have come to enjoy. For example, we collect data ranging from mouse clicks, site navigation, and links followed to time spent on a page, location data, and much more. This proprietary neural network allows us to take the data and customize what we present for the user to interact with.

This approach to a customized user experience does not come without its fair share of ethical concerns. Since the user does not have access or direct knowledge of the processes in the neural network it is possible for the user to inadvertently self-quarantine based on characteristics the algorithm might determine that the user finds preferential. This could include isolation of content from racial, political, or ethnic characteristic groups or individuals that, had the user been aware, might not consciously have chosen to isolate.

This approach to personalization is also not without legal concerns specifically from the GDPR enacted by the European Union. Let’s look at the portions of the GDPR that would impact our personalized user experience. First, we should consider transparency, obviously with how neural networks function the hidden layers would not be transparent to users especially in the sense that users won’t know exactly how their data is being used. Next, we can consider the principle of purpose limitation in which the data being gathered should be for a specific use and should not be archived.

The European commission’s policy is that data must be held for as little time as possible to process the data, that there must be a timetable for permanent data erasure, and that personal data is secured appropriately (2016 O.J. (L119/1) 5.1). Data accuracy describes how we should be keeping the data up to date to the best of our abilities, it is in our best interest to do this in order to keep our users’ preferences for personalization as up to the minute as possible.

When considering data minimization, we should look at reducing our collection of non-necessary data from the user. This works into reducing the costs associated with data collection in that reducing collection reduces storage requirements and limits the amount of data that we need to encrypt to ensure the confidentiality and security of the users personally identifiable information.

Finally, accountability allows providers like SMC to ensure that we are abiding by the laws and regulations of the countries in which our users reside and ensuring the well-being of our users by recording the measures we take to ensure compliance. Naturally there should be penalties should a company not take action to ensure compliance. These penalties can include loss of certification, fines ranging from 10 million Euro to 2% of global revenue, and loss of the ability to operate within the borders of the EU (2016 O.J. (L119/1) 83.2).

There are some areas of concern for the possible legal implications from the GDPR regarding SMC’s use of neural networks for a classifier to personalize the user’s experience. The biggest one has already been briefly touched on and that is the concern about transparency. The way neural networks naturally operate necessitate the capacity for opacity in how the data is processed. There is also the concern of data minimalization coupled with storage limitations.

While it is true that we would like to minimize the amount of data necessary to provide a unique personalized user experience, there may come a time in the future in which our researchers discover a specific data point that is key to delivering what the customer wants which may require the use of historic data. Considering that the GDPR makes room for the archival of data being used for scientific purposes there is some room for interpretation as to what the certifiers of the GDPR may feel is within the best interest of the public.

For our purposes at SMC the option of not collecting data is not a possibility. Without the collection of data we could still provide a service but the neural network that gives SMC a unique and marketable approach to a personalized user experience would not be able to function. It is also outside of the scope of the very idea of social media to not collect data. For example, if a user uploads a simple profile picture this would fall under the category of personal data that would need to be collected by SMC and stored for use whenever a user viewed this particular user’s profile page. Now we could take that understanding and apply it to all of the data that is processed throughout the implementation of our social media service. Without the ability for our neural network to recognize the users’ preferences through the data it collects the user would grow disinterested in the service that is not made specifically for them. Without collecting data SMC would cease to exist entirely robbing our users of a service they enjoy.

Thankfully SMC is not alone in the market of machine learning and many others have come up with principles and solutions for preserving users’ privacy while employing neural networks. One of the leading models for maintaining privacy in machine learning is Privacy-Preserving in Machine Learning (PPML) (Bhatt 2022). PPML has four essential tenants that help to ensure privacy is maintained throughout the entire pipeline.

These tenants include Data privacy in training which requires that the data used for training has been thoroughly vetted for privacy concerns. Privacy in Input which requires the inputted data to be opaque from the developers and possible malicious actors. Privacy in Output is similar in that it requires data in the output layer to have privacy measures in place to reduce or eliminate unauthorized parties from gaining access to the data. Finally, there is Model privacy to ensure that malicious actors cannot gain access to the neural network or steal its design aspects.

The manner in which SMC has been collecting, processing, and storing user data for personalization purposes should comply with all aspects of the GDPR. That being said, SMC will continue to audit their processes to ensure compliance with the GDPR. Transparency of our neural network being the biggest concern we have come up with a plan for our users to enter into a menu that allows them to gain a deeper understanding of how their data is being used and to customize how the algorithm helps to customize their experience including the ability to eliminate certain types of data from being collected and the chronological longevity of their data in its use in the neural network. With this change we hope to bring a new sense of transparency to our users that would act as a standard throughout the industry as a whole resulting in a better well-being for consumers of social media everywhere.

References:

Byrne, J. H. (2016). Introduction to neurons and neuronal networks. In Neuroscience online.

Retrieved from http://neuroscience.uth.tmc.edu/s1/introduction.html.

*What is an input layer? - definition from Techopedia*. Techopedia.com. (n.d.). Retrieved January 26, 2023, from https://www.techopedia.com/definition/33262/input-layer-neural-networks

DeepAI. (2019, May 17). *Hidden layer*. DeepAI. Retrieved January 26, 2023, 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.

Ognjanovski, G. (2020, June 7). *Everything you need to know about neural networks and backpropagation - machine learning made easy...* Medium. Retrieved January 26, 2023, from https://towardsdatascience.com/everything-you-need-to-know-about-neural-networks-and-backpropagation-machine-learning-made-easy-e5285bc2be3a

REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL, 2016 O.J. (L119/1) 5.1

REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL, 2016 O.J. (L119/1) 83.2

Bhatt, D. (2022, February 10). *Privacy-preserving in Machine Learning (PPML)*. Analytics Vidhya. Retrieved January 27, 2023, from https://www.analyticsvidhya.com/blog/2022/02/privacy-preserving-in-machine-learning-ppml/