Matthew Muller

Professor Obafemi Balogun

CS-370: Current/Emerging Trends in CS

26 March 2023

**White Paper: GDPR Compliance**

Artificial neural networks, or ANNs, like the one used by our company seek to create a system that is capable of learning how to best perform a task or set of tasks by analyzing training examples. They do this by loosely emulating the nature of the human brain, a network of billions of densely interconnected simple processing nodes called neurons that work together to perceive, learn, and act. ANNs create a less complex version of this network in which numerous nodes (sometimes thousands or even millions) are organized in layers consisting of an input layer, a number of hidden layers, and an output layer. The input layer receives data which is then passed through the hidden layers, being transformed each step of the way, until it reaches the output layer. Each node will have a number known as a “weight” assigned to each of its incoming connections and as the node receives data items, now in the form of numbers, it multiplies each incoming number by its associated weight and then sums these values, resulting in a single number. If this number is above the threshold value that has been set, the node “fires,” sending this number along all of its outgoing connections (Hardesty). This process continues as the data objects are passed through the hidden layers until they have been classified upon reaching the output stage.

Neural networks are already being used in a variety of ways to aid in the personalization of the user experience. These networks take in data from essentially all forms of interaction that a user has with their device and then use this data to customize their experience so that they are targeted by content which is deemed most likely to interest them. Sites like Amazon or Google can take advantage of access to a user’s cookies to collect huge amounts of personal and behavioral data, such as location, demographics, purchase and search histories, and much more. This data is then fed through the algorithm to determine things such as which adds a user is shown and which search results are sponsored going forward. While this is often convenient to the user, such as when being shown an ad for something they actually want to buy, it also raises a number of ethical concerns.

The first of these concerns is that of the possible invasion of the user’s privacy. Oftentimes, users are not aware of the extent of data collection that companies are performing. On top of this, they are left in the dark about how their data is being stored, managed, or used. This limits their ability to make informed decisions about how much of their privacy that they want to give up. Another possible ethical issue with these systems is the “black box” nature of them, in which the inputs and outputs can be observed but the inner workings are complex enough that they are not easily understood. While the rules for each node’s operation may be known, the performance of the entire system can be impossible to decipher due to the massive number of interacting nodes. The abstraction that is caused by this complexity can lead to systems developing biases without the developers realizing. Because of this, it is essential that great care is taken in the training of these algorithms to avoid the formation of biases.

The General Data Protection Regulation, or GDPR, is a set of rules put in place by the EU with the goal of harmonizing data privacy laws across all EU members as well as increasing protection of individuals’ personal information. Naturally, this set of regulations have a large impact on AI-based personalization as this function is based on the intake of massive amounts of user data. One of the main concerns addressed by this regulation is the lack of transparency in regards to how data is being collected and used. Up until this point, users were left in the dark for the most part when it came to how companies were collecting data from them and what that data was being used for. The GDPR now makes it a requirement that every form of data collection must be justifiable in its collection and use, and that users are informed accordingly. These rules go hand-in-hand with the concept of purpose limitation, as companies are no longer able to simply collect masses of data for any use that they see fit. Under the GDPR, companies must steer clear of data that does not have a direct use and only collect data that has a specific, pre-stated use. Another concept of user data protection that was addressed by the GDPR is that of storage limitation. In the past, companies were able to store user data indefinitely, holding onto it for any possible uses that arise. These changes require that data is stored only in order to be used for its pre-specified purposes and prohibit the continued storage of user data for unstated future uses. One more major concept of data security that is addressed by the GDPR is that of accountability. Previously, due to the lack of regulation, companies were able to do essentially whatever they wanted when it came to data collection/use. However, these changes make it so that anyone who collects user data is held responsible for following all principles of the GDPR and be penalized if not.

Since our company utilizes neural networks in order to provide personalization to our users, GDPR compliance is extremely important to us. In order to address any possible legal concerns of our company’s use of a neural network for personalization, we must take steps to ensure that we are compliant with all GDPR regulations. To begin, we must receive consent from users to collect data, which will be done through the implementation of opt-in checkboxes and disclosure of terms of data collection and use. Under the GDPR, double opt-ins are mandatory so we must ensure that this is being done before we begin to track a user’s data. While it is possible for our company to operate without collecting user data, it would greatly decrease the quality of the user experience and would therefore damage our business irrevocably. Much of our social networking experience is based upon anticipating the needs of our users and making recommendations based on this information. Without the collection of data from a user, providing this level of personalization is impossible.

There are many new efforts underway in pursuit of privacy under the GDPR. For example, “Data processing in untrusted environments - such as public cloud - and multiparty data sharing and analytics have become foundational to an organization’s success” (Gartner). This ensures that data collection and storage is being done in a transparent manner, with nothing being hidden from law or users. Another increasing trend in data privacy is data localisation. As more and more countries put data privacy laws in place, businesses have to put extra care into ensuring compliance with the laws of all involved countries if their data is being internationally used. “According to research, by 2024, 75% of the world’s population will have privacy laws protecting personal data” (Meghmala). It is essential that our company have employees dedicated to ensuring that our practices comply with all applicable laws as these legislations come into effect. Combined with our double opt-in for user data collection, educating ourselves on the GDPR and keeping up with new legislation will allow our company to achieve a reasonable level of personalization while maintaining legal and ethical responsibility.

**Works Cited**

Burgess, Matt. “What Is GDPR? the Summary Guide to GDPR Compliance in the UK.” WIRED

UK, 24 Mar. 2020, https://www.wired.co.uk/article/what-is-gdpr-uk-eu-legislation-compliance-summary-fines-2018.

Card, Dallas. “The ‘Black Box’ Metaphor in Machine Learning.” Medium, Medium, 14 July

2021,https://dallascard.medium.com/the-black-box-metaphor-in-machine-learning-4e57a3a1d2b0.

“Gartner Identifies Top Five Trends in Privacy through 2024.” Gartner,

https://www.gartner.com/en/newsroom/press-releases/2022-05-31-gartner-identifies-top-five-trends-in-privacy-through-2024.

“GDPR: How Does It Affect Social Media? — Riskoptics - Reciprocity.” Reciprocity, 22 Nov.

2021, https://reciprocity.com/blog/gdpr-how-does-it-affect-social-media/.

Hardesty, Larry. “Explained: Neural Networks.” MIT News | Massachusetts Institute of

Technology, https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414.

Meghmala. “Top Data Privacy Trends to Watch out for in 2023.” Analytics Insight, 21 Feb.

2023, https://www.analyticsinsight.net/top-data-privacy-trends-to-watch-out-for-in-2023/.