Privacy in the Age of AI: Navigating the Ethical Dimensions of Machine Learning

With the current exponential growth of computer processing capabilities increasing to the point where consumer graphics cards can easily train and run advanced machine learning models, there is increasing concern about the ethical factor of privacy in this “Age of AI”. With the growth of AI, there is a greater demand for data that leads to unethical data collection methods, and more advanced models can accurately infer sensitive information based on other seemingly unrelated data. This need for people to maintain their privacy requires AI creators to ethically collect and use data, and governments to regulate and enforce this.

To use machine learning to create an artificial intelligence (AI) model, an incredibly large amount of data is needed to train models. For example, GPT-4 required a training data set with an estimated 1.8 trillion parameters. The ethical concern is in the methods used to collect this data, especially when the data is about people. It is extremely difficult to get so much data, especially for large projects associated with data about people, so automatic data collection and scraping needs to be conducted. Automatic data scraping is where data searching is conducted with particular relevant parameters set by the person training a model, however there can be little to no regulation of where that data is coming from, meaning sensitive data could be collected from personal records. This is one of the only viable methods for training models, and can be very intrusive to privacy and is conducted without consent from the people. While unethical, this is technically legal, which makes it a significant issue to privacy and will only get worse with the progression of larger AI with the exponential increase of processing power. As well as this, the large amount of data collected about people can be used to conduct attacks that can infer other sensitive information without ever being in contact with it.

Machine learning models can be trained to conduct AI inference attacks which are difficult to prevent and are a great threat to privacy. Machine learning opens up this whole new angle of attacks, where AI can be trained to accurately infer sensitive information about people based on other seemingly unrelated data. For example, if information about a person’s family, status and income is inputted into a trained model, it can accurately output the possibility for their loan to default which can be used maliciously by an attacker. It is difficult for people to maintain privacy against inference attacks as they are very difficult to anticipate or prevent without drastic measures.

However, model creators can add algorithms to models to prevent malicious use from others, for example ChatGPT's censorship filter which prevents things such as users being able to research how to conduct criminal activities. Due to this, creators of AI models must be held responsible to make sure that malicious users cannot use machine learning to train an AI in such a way.

In order for people to maintain their privacy in this new Age of AI, there is a moral responsibility for AI model creators to gather data ethically while using it for legitimate purposes, as well as making sure others cannot fine tune their models for malicious use. In addition to this, governments must have regulations for the use of AI, and also enforce them. Many governments such as the Australian government only have a set of eight AI ethics principles that are encouraged but unenforced, which is an issue as this is the only viable way for everyday people to be protected from privacy intrusion by unethical data gathering or attacks. An approach more similar to the European Union with their Artificial Intelligence Act where ethical requirements are mandated and regulated needs to reach a global scale to fully make sure AI models need to be developed ethically.

With this new rise of machine learning capabilities, privacy has become increasingly difficult to maintain with greater possibilities of having personal data scraped or predicted by AI. This ethical dilemma that is created requires ethical practices for machine learning to be enforced by law, otherwise this issue will only worsen over time.

References

https://www.w3.org/TR/webmachinelearning-ethics/
https://www.prolific.co/blog/ai-data-scraping-ethics-and-data-quality-challenges
https://portswigger.net/daily-swig/inference-attacks-how-much-information-can-machine-learning-models-leak
https://medium.com/predict/gpt-4-everything-you-want-to-know-about-openais-new-ai-model-a5977b42e495
https://pub.towardsai.net/inference-attacks-the-sql-injection-of-the-future-ba6daa563682
https://towardsai.net/p/machine-learning/so-what-are-inference-attacks
https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles
https://www.michalsons.com/blog/membership-inference-attacks-a-new-ai-security-risk/64440