Zachary Nicholas

March 28, 2024

CS 370 Current and Emerging Trends in CS

Project One

1. **Explain the basics of neural networks and how they work** by addressing the following:
   1. Provide a brief explanation of how neural networks work. How do the input layer, hidden layer, and output layer interact to classify objects? Consider the fact that your target audience may have limited technical knowledge.

The basics of a Neural Network (NN) are the different layers that they consist of, which are the input, hidden, and output layers to make these more understandable to our daily lives the input layer would be sort of like when we take something in from one of our five senses, The hidden layer would be sort of like our intuition of how we think something should or would be like which then gets checked by our brain such as smelling something and thinking that it will taste bad and then we taste it and we confirm if the smell is representative of the taste it has, then the output of this is telling how the smell and taste are associated. That’s a very simple rendition of how they work but they are also a lot more in-depth in terms of how these processes are completed with an NN, For instance, A NN isn’t as limited as my example process there can often be many different individual parts in a layer like there can be more than just 1 to 2 of an input layer to relate it to my example it's like using more senses except taste such as sight, feeling, smell and potentially hearing based off the item, The same goes for the hidden layers as these can be almost infinite the only thing that matters is the amount of data that is fed to the NN as not feeding it enough could just waste time and energy of having them active when they aren’t active and often comes to a balancing act to make sure you have enough data as well as enough hidden layers to process the data, to relate this to my brain example would be like to constantly keep analyzing the data until a solid answer is reached. The output layers of the NN are similar to the Hidden layers in the sense that it depends on the question that is being asked of it and how complex the answer should be the simplest answer would be a yes or no, or a true or false. To relate it to my example of the human brain, The question “Does it taste like it smells?” with the answer that would be required of the brain being a yes or no, another question that would require more than one output would be something like “Does the Banana seem ripe? “Approximately how many days does it have till it goes bad?” this answer would require at least two different outputs one being a yes or no and the other requiring a bit more complex.

1. **Evaluate how neural networks are used to create personalization** by addressing the following:
   1. How are neural networks utilized to aid in the personalization of the user experience?

One way that almost everyone has experience with is either the “for you” page on TikTok or the recommended videos on YouTube either through the actual recommendation feature or through the mix feature that will look at the specific things that you either watch a lot or things that you use the like feature on. What this does is compile a simple user profile to help assign videos that the current user would like based on videos the user has watched previously, the current trending videos, and videos by creators you’ve watched before. For instance, in my use case for YouTube, I mainly watch a lot of music reaction channels, Tech channels, and various video games and anime channels which makes my user experience and videos I get recommended wildly different from someone who only says watches Tech channels like LTT (Linus Tech Tips) for more consumer tech or someone like Level1Techs who offers a mix of end-user equipment for stuff like a home lab or things for a small company.

How I believe that a company could find this useful is the fact that a user would spend more time with an app that tailors itself to what the user wants and keeps consuming like in my example of my search history and the things that I am recommended by YouTube to watch.

* 1. What ethical concerns can this raise? Consider hidden biases that may arise in using a “black box” classification system, where the algorithms are unknown to the user.

Like in my example I think a hidden bias is that the algorithm that YouTube uses doesn’t recommend a lot of different races even though I often watch all different sorts of people but it seems to only stick me with white creators (Which I am fine with for the most part but there are certain channel I rarely get recommended but would like to see more of them such as Marques Brownlee who does a lot of apple reviews which is something that is also lacking from my recommendations). Other examples of an AI that gets fed certain data and produces results that aren’t intended and cause a bias against a group would be the Amazon AI that after looking at 10 years of applications of workers at Amazon found that women were being submitted far less than their male counterparts so with the data it was given it assumed that women were not the people that were trying to be hired. Another big concern for me is the profile that something like a YouTube or a Facebook makes on me and how it can be used to either sell to advertisers or make it to the point where I don’t leave the app for an extended period because they keep serving me interesting content. Also with YouTube AI, we don’t know specifically the data it collects and how it uses that data such as my earlier point of selling our data to Ad companies or even exactly how the recommendation feature works.

1. **Analyze how portions of the GDPR affect personalization** by addressing the following:
   1. Summarize the portions of the GDPR that affect personalization. Be sure to consider *at least four* of the following in your answer: transparency, purpose limitation, data minimization, accuracy, storage limitation, confidentiality, and accountability.

With GDPR it didn’t so much affect personalization it more is trying to balance what information the user is asked to give and how it could be used through transparency. With that however, the increase of users worrying about their data has made it a necessity to be more transparent with the users of your product and how that data will be used especially if the data could be either leaked or otherwise sold to an Ad agency. Another issue that has become more apparent as of late is that often unless the AI is being tailored to the user it will feel very random for the things they are suggested in my example earlier on YouTube you can see this if you log out of your account and look at the recommended page it will be very random compare to what might interest you (in other words the push for confidentiality has affected the overall accuracy of a system). Another problem with certain apps is their overreaching into data they don’t need access to that could be used against you this has made companies all around more transparent and require less information from the individual user which also lowers the risk of being accountable if the data that they collect were to leak or get out another way.

1. **Assess how the GDPR is affecting the company’s practices** by addressing the following:
   1. What specific legal concerns may arise from your company’s use of neural networks as a classifier to personalize the user experience?

As I had started to say above, companies are currently being pressured by both end users of their software but also the countries that represent them to prevent the unwanted leakage of personal data such as the hack that happened to Facebook back in August of 2019 where over 530 million users that had their data taken and later posted in a public database until April 2021, with some of that data including phone number, addresses, email address, and their full names as well as other things that Facebook had collected, This had an irreversible effect to the company’s overall community support and only got worse as a whistleblower in October 2021 came out and said that Facebook often chose to focus on profits instead of the safety of their data. This is a serious issue that a company could run into regardless of size and the measures that they take to protect that sensitivity that would be collected to help recommend more content to its users.

* 1. Is not collecting data a possibility for the company’s business model? Defend your answer.

In this business model, it is required to at least collect some data to serve better content to an end user but it should come with certain checks to make sure we aren’t storing too much sensitive data, and if we are make sure that it is well protected. In my two cases, I brought up YouTube and Facebook with one using data that isn’t all too sensitive and still getting the job done of recommending certain content that might interest the users, and Facebook arguably collects too much data and didn’t make better considerations to make sure that the data was protected and couldn’t be stolen by a nefarious actor. I think that for our company to succeed using a similar model that YouTube does where it takes a lot of information that isn’t too sensitive as compared to all of the data that Facebook collects and most likely sells to Ad companies (Allegedly).

1. **Propose adaptations to the company’s practices to act in compliance with the GDPR** by addressing the following:
   1. What are the current trends (best practices) in artificial intelligence and machine learning aimed at preserving privacy?

Most of the trends and best practices I could find relate directly back to GDPR and the overall focus on using only purpose-driven and limited machines to prevent the unwanted leakage of data these could be done in two different types one where all of the data is stored in one location and secure from any tampering by restricting access to it unless necessary these are often thought of as data enclaves the other method being federated learning where the data is physically separated into different sites and then ran from there. Another thing that is common practice is to use Data Minimization and storage limitation to make sure that the data that is being collected from the user is necessary and not irrelevant with that also comes trying to anonymize the data that is being collected therefore anyone that could get their hands on the data wouldn’t have all of the data from that person on the bits and pieces of it and finally on this point limiting the users that have access to this data is crucial and can further prevent a breach or a heist of the data from your storage. And lastly being transparent with your users and obtaining their consent to store and use this data is another huge part of making sure that the data is okay to be there and that the end user of the product finds it to be useful instead of a generalized recommendation that they would have before giving this data.

* 1. What changes to the way the company collects, stores, and employs user data do you propose to comply with GDPR? Defend existing practices where applicable.

As I stated above, the best thing to do in this situation is to be fully transparent about the data being collected and how it will be used and obtain consent from the user that this data will be taken to provide them with a better overall user experience. Another part of it is making sure that the data can’t be very easily traced back to the original user and used against them in any way. I think that this data is okay to gather from the user as long as it is encrypted and anonymized so that in the event of a data breach or leak it doesn’t become a major problem for them and we can continue to collect the revenue by suggesting the not only content they are interested in but also in ads that will interest them as well.

References:

Arcserve. (2023, December 20). *7 most infamous cloud security breaches*. https://www.arcserve.com/blog/7-most-infamous-cloud-security-breaches

*Deep learning in Personalization*. Recosense Labs Inc. (2021, September 28). https://www.recosenselabs.com/blog/deep-learning-in-personalization

Freischlag, C. (2021, August 7). *Using machine learning to personalize user experience*. Medium. https://towardsdatascience.com/using-machine-learning-to-personalize-user-experience-f5b6abd65602

Geyser, W. (2024, March 18). *How does the YouTube algorithm work: A peek into its changes in 2024*. Influencer Marketing Hub. https://influencermarketinghub.com/how-does-the-youtube-algorithm-work/#toc-0

GfG. (2024, January 3). *What is a neural network?*. GeeksforGeeks. https://www.geeksforgeeks.org/neural-networks-a-beginners-guide/

IBM developer. (n.d.). https://developer.ibm.com/articles/l-neural/

Lavanchy, M. (2023, August 8). *Amazon’s sexist hiring algorithm could still be better than a human*. IMD business school for management and leadership courses. https://www.imd.org/research-knowledge/digital/articles/amazons-sexist-hiring-algorithm-could-still-be-better-than-a-human/

MacCarthy, M., Darrell M. West, J. R. A., Zhang, B., Daniel S. Hamilton, J. R., O’Hanlon, M. E., Clifford Winston, J. Y. S. D., Jacob Larson, J. S. D., & Darrell M. West, M. K. (2023, June 27). *Protecting privacy in an AI-Driven World*. Brookings. https://www.brookings.edu/articles/protecting-privacy-in-an-ai-driven-world/

Monetate. (2023, September 25). *How AI improves customer experience personalization*. https://monetate.com/resources/how-ai-is-transforming-customer-experience-through-personalization-strategies/

*OWASP AI Security and Privacy Guide*. OWASP AI Security and Privacy Guide | OWASP Foundation. (n.d.). https://owasp.org/www-project-ai-security-and-privacy-guide/

O’Reilly-Shah, V. N., Gentry, K. R., Walters, A. M., Zivot, J., Anderson, C. T., & Tighe, P. J. (2020, December). *Bias and ethical considerations in machine learning and the automation of Perioperative Risk Assessment*. British journal of anaesthesia. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7442146/

Schindler, N. (2020, July 9). *The implications of GDPR on personalization*. Dynamic Yield. https://www.dynamicyield.com/blog/gdpr-implications-personalization/

Shahani, A., & Schneider, A. (2019, July 24). *FTC to hold Facebook CEO Mark Zuckerberg liable for any future privacy violations*. NPR. https://www.npr.org/2019/07/24/741282397/facebook-to-pay-5-billion-to-settle-ftc-privacy-case

Team, T. A. (2021, May 28). *How is YouTube using AI to recommend videos?* https://towardsai.net/p/l/how-is-youtube-using-ai-to-recommend-videos

Theriot, K. (2023, August 11). *Council post: Marketing in the age of privacy: Balancing personalization and data protection*. Forbes. https://www.forbes.com/sites/forbesagencycouncil/2023/08/09/marketing-in-the-age-of-privacy-balancing-personalization-and-data-protection/?sh=2f86c95d4d09

What is a neural network? - artificial neural network explained - AWS. (n.d.). https://aws.amazon.com/what-is/neural-network/

Yang, B., Wei, L., & Pu, Z. (2020, November 19). *Measuring and improving user experience through artificial intelligence-aided design*. Frontiers in psychology. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7710987/