Machine Learning - Natural Language Processing (Unit 10)

Overview

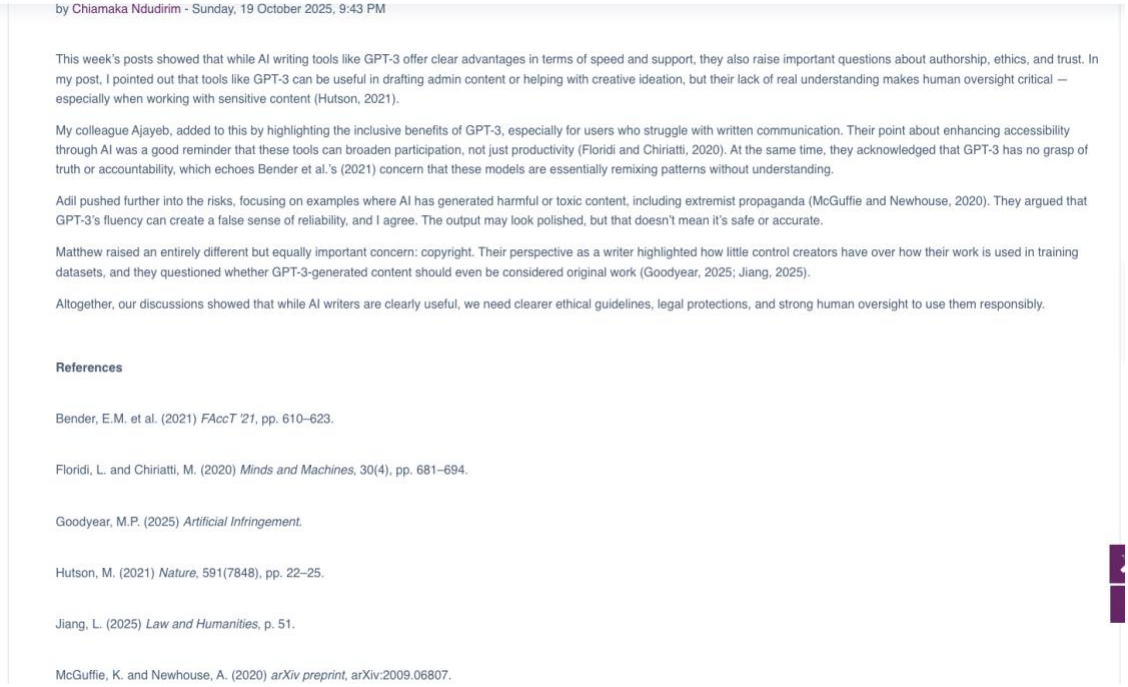
This week covered Natural Language Processing (NLP) and how it has evolved from rule-based systems to powerful transformer models like BERT, GPT, and T5 (Vaswani et al., 2017). We looked at how these models are applied in tasks like translation, chatbots, and summarisation, and how they're evaluated using metrics like BLEU and ROUGE.

What I Have Learned

I gained a clearer understanding of how transformer models work and the role of self-supervised learning in improving them. The evaluation metrics helped me see how NLP outputs are measured, and the link to MLOps showed how models are maintained in real-world use. Overall, it tied technical concepts to practical applications.

Collaboration Discussion: Summary

This week's e-portfolio reflects on our discussion of Hutson's (2021) article on AI writers like GPT-3. My summary post brought together peer views on inclusion, misinformation, and copyright. While the tech has potential, Hutson highlights the need for oversight. The discussion reinforced that ethical use must be guided by clear regulation and accountability. A screenshot of my full post is below.



by Chiamaka Ndudirim - Sunday, 19 October 2025, 9:43 PM

This week's posts showed that while AI writing tools like GPT-3 offer clear advantages in terms of speed and support, they also raise important questions about authorship, ethics, and trust. In my post, I pointed out that tools like GPT-3 can be useful in drafting admin content or helping with creative ideation, but their lack of real understanding makes human oversight critical — especially when working with sensitive content (Hutson, 2021).

My colleague Ajayeb, added to this by highlighting the inclusive benefits of GPT-3, especially for users who struggle with written communication. Their point about enhancing accessibility through AI was a good reminder that these tools can broaden participation, not just productivity (Floridi and Chiriatti, 2020). At the same time, they acknowledged that GPT-3 has no grasp of truth or accountability, which echoes Bender et al.'s (2021) concern that these models are essentially remixing patterns without understanding.

Adil pushed further into the risks, focusing on examples where AI has generated harmful or toxic content, including extremist propaganda (McGuffie and Newhouse, 2020). They argued that GPT-3's fluency can create a false sense of reliability, and I agree. The output may look polished, but that doesn't mean it's safe or accurate.

Matthew raised an entirely different but equally important concern: copyright. Their perspective as a writer highlighted how little control creators have over how their work is used in training datasets, and they questioned whether GPT-3-generated content should even be considered original work (Goodyear, 2025; Jiang, 2025).

Altogether, our discussions showed that while AI writers are clearly useful, we need clearer ethical guidelines, legal protections, and strong human oversight to use them responsibly.

References

Bender, E.M. et al. (2021) *FAccT '21*, pp. 610–623.

Floridi, L. and Chiriatti, M. (2020) *Minds and Machines*, 30(4), pp. 681–694.

Goodyear, M.P. (2025) *Artificial Infringement.*

Hutson, M. (2021) *Nature*, 591(7848), pp. 22–25.

Jiang, L. (2025) *Law and Humanities*, p. 51.

McGuffie, K. and Newhouse, A. (2020) *arXiv preprint*, arXiv:2009.06807.

References

Hutson, M. (2021) ‘The language machines’, *Nature*, 591(7848), pp. 22–25.

Vaswani, A. et al. (2017) ‘Attention is all you need’, *Advances in Neural Information Processing Systems*, 30, pp. 5998–6008.