Toward More Reliable Learning Models through Human-Inspired Learning

Nicki Barari[1], **Christopher J. MacLellan**[2], **Edward Kim** [1]

[1]Drexel University, Philadelphia, PA, USA
[2]Georgia Institute of Technology, Atlanta, GA, USA
nicki.barari@drexel.edu, cmaclell@gatech.edu, ek826@drexel.edu

Abstract

In recent years there has been an incredible progress in the fields of machine learning and computer vision, which has led to remarkable achievements, set new standards, and demonstrated the potential to surpass human capabilities in certain tasks such as object detection and classification (Krizhevsky, Sutskever, and Hinton 2012; He et al. 2016).

These breakthroughs are impressive; however, machine learning models still fall behind human capabilities in tasks that require deep understanding and adaptability. Issues such as catastrophic forgetting (McCloskey and Cohen 1989), high computational cost, and the need for adaptive learning architectures highlight the gap between current machine learning capabilities and the dynamic nature of human knowledge acquisition. This work emphasizes the necessity of integrating human-inspired learning approaches to bridge this gap and aims to provide machines with learning mechanisms that are more detailed, context-aware, and adaptable.

This work identifies three essential features for achieving a more human-like learning model: 1) Continual, lifelong learning, which refers to the ability of models to continually learn and adapt to new tasks, over time, without forgetting previously acquired knowledge (Parisi et al. 2019); 2) Transfer learning, which refers to the ability of learning models to apply knowledge learned from one task or domain to improve the performance on a different, but related, task (Yosinski et al. 2014); and 3) Robustness in real-world applications, which refers to the ability of a model to perform reliably and accurately in a dynamic and unpredictable environment (Goodfellow, Shlens, and Szegedy 2014). These key features help to emulate the efficiency, flexibility, and robustness of human knowledge acquisition.

By exploring key concepts and methodologies over time, this work aims to identify ongoing challenges, and highlight potential directions for future research in the pursuit of human-like learning. The journey towards creating machine learning models that are more human-like is challenging, and requires a deeper understanding of human cognition and learning processes, and as the field of machine learning continues to evolve, human-inspired approaches hold the key to unlocking more advanced, reliable, and adaptable learning systems.

References

Goodfellow, I. J.; Shlens, J.; and Szegedy, C. 2014. Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572*.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.

Krizhevsky, A.; Sutskever, I.; and Hinton, G. E. 2012. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.

McCloskey, M.; and Cohen, N. J. 1989. Catastrophic interference in connectionist networks: The sequential learning problem. In *Psychology of learning and motivation*, volume 24, 109–165. Elsevier.

Parisi, G. I.; Kemker, R.; Part, J. L.; Kanan, C.; and Wermter, S. 2019. Continual lifelong learning with neural networks: A review. *Neural networks*, 113: 54–71.

Yosinski, J.; Clune, J.; Bengio, Y.; and Lipson, H. 2014. How transferable are features in deep neural networks? *Advances in neural information processing systems*, 27.