ML is NOT glorified curve-fitting.

Jimut Bahan Pal
22D1594

February 9, 2023

Machine Learning is the superset of Deep Learning and a subset of Artificial Intelligence algorithms [1] that can be used for various tasks, from weather forecasting, sentiment classification, image segmentation etc. Machine Learning is taking over the world, and some algorithms are so complex that they can be compared with a black box. Statistical Machine learning algorithms are transparent and explainable, but their capabilities are limited. The current form of Machine Learning, i.e., Deep Learning, involves building complex models which are essentially black boxes are taking over all the intricate and unconventional tasks, like image classification, text generation, text completion, etc., which were not possible with statistical machine learning with that level of accuracy and precision.

My research lies in the field of deep learning, and computer vision specifically applied to medical settings in the domain of image segmentation. In medical sectors, doctors are pretty skeptical of using deep learning models, which are not interpretable [2], since the medical domain involves high-stake decisions which can cost lives. Suppose a deep learning model is used to predict whether a specific type of tissue in the image is cancerous or not; if the deep learning model cannot predict cancer when the patient has cancer, then the patient might go untreated. In the other case, the patient might sit for an operation while he doesn't have cancer. Both cause inconvenience to the patient and doctors; hence, our method will not wholly replace doctors but will aid doctors in coming to logical and plausible explanations. This cannot be done with curve fitting since curve fitting is interpretable; the maximum one can interpret deep learning models in the domain of computer vision is

where the models are looking into and what parts of the image are actually helping it to come to a conclusion [3]. The deep learning models are very complex; they can model functions in an abstract way; for example, researchers have found that the initial layers of the deep learning models learn texture-based filters, and the filters become more holistic, i.e., class-based, when the model reaches the final layer. Some researchers believe that the models are actually learning manifolds for generating data with the help of generative networks, but those are abstract representations, and no one is quite sure what is actually happening inside the models.

The domain of machine learning in the field of deep learning is still dormant, i.e., researchers are not interested in creating interpretable models but more complex models which achieve superior performance than interpretable models. The current form of interpretable models deals with a surrogate model which can interpret the deep learning models. Deep Learning is rarely given complete control in high stake tasks and is used with human supervision. Curve-fitting algorithms have limited performance, and the designer knows when things might mess up, but the deep learning algorithm messes things up confidently.

References

[1] Iqbal H Sarker. Machine learning: Algorithms, real-world applications and research directions. *SN computer science*, 2(3):160, 2021.

[2] Amitojdeep Singh, Sourya Sengupta, and Vasudevan Lakshminarayanan. Explainable deep learning models in medical image analysis. *Journal of Imaging*, 6(6), 2020.

[3] Cynthia Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature machine intelligence*, 1(5):206–215, 2019.