Bio-inspired robust machine learning

[1,2,3] Ilona Kulikovskikh, [2] Tomislav Šmuc

[1] Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia
Centre of Research Excellence for Data Science and Advanced Cooperative Systems
[2] Division of Electronics, Ruđer Bošković Institute, Zagreb, Croatia
[3] Samara National Research University, Moskovskoe Shosse 34, Samara, Russia

E-mail: ilona@irb.hr, smuc@irb.hr

Modern machine learning algorithms can successfully tackle tough and complicated problems by taking patterns buried inside datasets to build a model with remarkable predictive capabilities. However, the machine learning models fueling innovations in a variety of applications from large-scale genomic sequencing and medicine to automated driving and robotics still pose serious challenges that make it difficult to fully trust and adopt them. A lack of intelligence in these models leads to an inability to robustly differentiate between noise and subtle, but significant variation in data. If the noise in data includes intentionally small carefully crafted perturbations, which are used to generate so-called adversarial examples, the model may become vulnerable and misclassify them with high confidence. It expectedly sets up the psychological roadblocks to the widespread adoption of machine learning models in different domains.

Previous studies suggested various attacking strategies to fool the model with adversarial examples as well as defenses to resist them, but that has not solved the problem completely. What is not specifically undertaken in the studies mentioned above is whether the models inspired by biological evolution may result in better robustness to adversarial perturbations. This is the research question raised in this study. The fundamental aspects of biological intelligence, such as self-healing, evolution, and learning make biological organisms successful to survive in unknown and changing environments. The stability and adaptability of biological systems strengthen the motivation for replicating the mechanisms of natural evolution in an attempt to create the models with characteristics comparable to those of biological systems.

This paper seeks to refine the discourse on robustness to adversarial noise with the bio-inspired machine learning based on the generalized Verhulst population growth equation. While using the common approach to penalizing the large input gradients in an optimization procedure, which are more likely to be utilized for generating adversarial examples, we optimize hyperparameters of the Verhulst equation to penalize the degree to which imperceptible changes in gradients may influence

prediction results. We refer to them as *the lower and upper levels of visibility* as they limit the gradients to be near zero so that any small magnitude perturbation hidden in the gradients is “visible” and, then, has no influence on prediction results.

We analyzed the impact of the hyperparameters on the robustness properties empirically in order to support the results of theoretical outcomes. First, we justified the improvement in robustness and performance of the bio-inspired model over the simple model on the synthetic linearly separable dataset with different proportions of noisy labels. Then, we assessed the viability of the bio-inspired model with the hyperparameters in the more realistic setting on 11 experimental datasets, which are freely available from UCI Machine Learning repository. The chosen datasets are normally not linearly separable but are indicative of the behavior of the hyperparameters and their influence on prediction results. Finally, we applied the bio-inspired optimization technique to a more complicated model on a large dataset for multiclass image classification. The results of computational experiments showed consistency and confirm the viability of bio-inspired machine learning models. We concluded that they may contribute to better robustness to adversarial noise by penalizing the degree to which subtle changes in gradients impact prediction.

References

[1] Kulikovskikh I, Prokhorov S, Lipić T, Legović T, Šmuc T. BioGD: Bio-inspired robust gradient descent. 2019. PLoS ONE 14(7): e0219004. https://doi.org/10.1371/journal.pone.0219004

[2] Kulikovskikh I, Prokhorov S, Legović T, Šmuc T. Growing descent of stochastic gradient with the generalized logistic map. V International Conference on Information Technology and Nanotechnology (ITNT). 2019: 338-344.