Generalizing the Information-theoretic Characterization of Channels That Compute

Aman Chawla

March 22, 2024

Abstract

In this note, the authors discuss a perceptron as a channel that computes.

In the prescient work [1], Berger and Levy introduced a mathematical model of a neuron as a multiple-access, partially degraded broadcast channel. They took steps towards its Shannon capacity characterization, along with a discussion of the data processing it performs. We know that the perceptron algorithm closely follows a neuron model. Thus Berger and Levy in a sense took steps toward the "capacity" of the perceptron. The perceptron is only one of many machine learning (ML) algorithms. These algorithms, once their weights are settled, have been "trained" to compute a certain function of their input and present it at the output.

A trained ML algorithm is not only a function approximator but also a communication channel. In this sense it is a channel that computes. By extending the Berger-Levy paradigm to various ML algorithms, we can thereby characterize their usefulness as data transmission devices. With this joint action of the transmission of information and the performance of computation and compatible combinations of these two, these devices can perhaps find novel use in hardware such as AI chips. If two AI chip submodules need to be interlinked, such a device might offer efficiencies not available with either a channel or a ML algorithm stand-alone.

References

[1] Toby Berger and William B Levy. A mathematical theory of energy efficient neural computation and communication. *IEEE Transactions on Information Theory*, 56(2):852–874, 2010.