Data Shapley: Equitable Valuation of Data for Machine Learning

Connor Levenson

August 13, 2020

This paper by Amirata Ghorbani and Dr. James Zou can be found here.

In machine learning there is the acronym GIGO - which stands for "garbage in, garbage out." The main idea is that if you are going to use data to understand the dynamics of some system (and hopefully make some predictions), then the data used in the training process must be a good representation of the population as a whole. Even with the most sophisticated models, if the data you are extracting information from is "garbage," then the results of your model will be useless. Point being: data quality is very important. This paper proposes multiple algorithms for approximating the "Shapley" value of a data point in a supervised machine learning setting. The Shapley value derives from cooperative game theory, and can be thought of as the value that a certain member adds to a coalition (example: what value does a certain country add to an alliance). In this case, the coalition is the entire training set, and each member is an individual training point. The idea is to compare the performance of the coalition with different permutations of members, and see how that performance of the coalition changes with different members. We use this information to make insights about what each member adds to the whole group.

The paper proposes two algorithms: TMC-Shap (Truncated Monte Carlo Shapley) and G-Shap (Gradient Shapley). The latter makes a beautiful connection between the Shapley value and Stochastic Gradient descent, leading to an efficient way to approximate the Data Shapley values for machine learning models that use SGD for optimization.

Connor is interested in interpretability of deep learning models. He says that "if we are going to deploy

models that make decisions that risk the lives of human beings, we must understand how the model came to its prediction." Gradient Shapley is an efficient way to measure the quality of individual training points of deep learning models. One important application that the paper shows is that the Data Shapley value can be used to identify mis-labeled data - something that can be harmful to the generalizability of a model. Due to this, Connor is interested in more applications of the Data Shapley value (for example, what if we could understand why data is [not] valuable to our model).

Location and Time

Zoom @ 5:30pm - 6:30pm (PST), a zoom invite will be sent out around 4:15pm

Contact

Website
Email: numerical.scouts.ucsb. at gmail.com