

## Data Depth as a Risk

Data depths are score functions that quantify in an unsupervised fashion how central is a point inside a distribution, with numerous applications such as anomaly detection, multivariate or functional data analysis, arising across various fields. The halfspace depth was the first depth to aim at generalising the notion of quantile beyond the univariate case.

Among the existing variety of depth definitions, it remains one of the most used notions of data depth. Taking a different angle from the quantile point of view, we show that the halfspace depth can also be regarded as the minimum loss of a set of classifiers for a specific labelling of the points. By changing the loss or the set of classifiers considered, this new angle naturally leads to a family of "loss depths", extending to well-studied classifiers such as, e.g., SVM or logistic regression, among others. This framework directly inherits computational efficiency of existing machine learning algorithms as well as their fast statistical convergence rates, and opens the data depth realm to the high-dimensional setting. Furthermore, the new loss depths highlight a connection between the dataset and the right amount of complexity or simplicity of the classifiers. The simplicity of classifiers as well as the interpretation as a risk makes our new kind of data depth easy to explain, yet efficient for anomaly detection, as is shown by experiments.