Abstract:

Universal learning is considered from an information theoretic point of view following the universal prediction approach pursued in the 90's by F&Merhav. Interestingly, the extension to learning is not straight-forward. In previous works we considered on-line learning and supervised learning in a stochastic setting. Yet, the most challenging case is batch learning where prediction is done on a test sample once the entire training data is observed, in the individual setting where the features and labels, both training and test, are specific individual quantities. This work provides schemes that for any individual data compete with a "genie" (or a reference) that knows the true test label. It suggests design criteria and derive the corresponding universal learning schemes. The main proposed scheme is termed Predictive Normalized Maximum Likelihood (pNML). As demonstrated, pNML learning and its variations provide robust and "stable" learning solutions that outperform the current leading approach based on Empirical Risk Minimization (ERM). Furthermore, the pNML construction provides a pointwise indication for the learnability that measures the uncertainty in learning the specific test challenge with the given training examples; hence the learner knows when it does not know. The improved performance of the pNML, the induced learnability measure and its utilization are demonstrated in several learning problems including deep neural networks models. Joint work with Yaniv Fogel and Koby Bibas.