Contribution ID: **16**   Type: **not specified**

Fisher information metrics for binary classifier evaluation and training

Wednesday, 11 April 2018 09:05 (20 minutes)

Different evaluation metrics for binary classifiers are appropriate to different scientific domains and even to different problems within the same domain. This presentation focuses on the optimisation of event selection to minimise statistical errors in HEP parameter estimation, a problem that is best analysed in terms of the maximisation of Fisher information about the measured parameters. After describing a general formalism to derive evaluation metrics based on Fisher information, three more specific metrics are introduced for the measurements of signal cross sections in counting experiments (FIP1) or distribution fits (FIP2) and for the measurements of other parameters from distribution fits (FIP3). The FIP2 metric is particularly interesting because it can be derived from any ROC curve, provided that prevalence is also known. In addition to its relation to measurement errors when used as an evaluation criterion (which makes it more interesting that the ROC AUC), a further advantage of the FIP2 metric is that it can also be directly used for training decision trees (instead of the Shannon entropy or Gini coefficient). Preliminary results based on the Python sklearn framework are presented. The problem of overtraining for these classifiers is also briefly discussed, in terms of the difference of the FIP2 metric on the validation and training set, and of their difference from the theoretical limit. Finally, the expected Fisher information gain from completely random branch splits in the decision tree and its possible relevance in reducing overtraining is analysed.

Intended contribution length

20 minutes

Primary author: VALASSI, Andrea (CERN)

Presenter: VALASSI, Andrea (CERN)

Session Classification: Session 5