Contribution ID: **78** Type: **Parallel talk**

Semi-Supervised Anomaly Detection - Towards Model-Independent Searches of New Physics

Tuesday, 6 September 2011 16:35 (25 minutes)

Most classification algorithms used in high energy physics fall under the category of supervised machine learning. Such methods require a training set containing both signal and background events and are prone to classification errors should this training data be systematically inaccurate for example due to the assumed MC model. To complement such model-dependent searches, we propose an algorithm based on anomaly detection techniques, which does not require a MC training sample for the signal data. We first model the MC background using multivariate mixtures of Gaussians. We then search for deviations from the background model by fitting to the observations a mixture of the background model and a number of additional Gaussians using a variant of the EM algorithm. This allows us to perform pattern recognition of any excess over the background. We show by comparison to neural networks that such a semi-supervised approach is a lot more robust against misspecification of the signal MC than supervised classification. In cases where there is an unexpected signal, a neural network fails to correctly identify it while anomaly detection does not suffer from such a limitation. On the other hand, when there are no systematic errors in the signal MC, both methods perform comparably. Due to its fully probabilistic nature, the anomaly detection model has a number of additional advantages as well. Firstly, the mixing proportion of the anomalous excess immediately gives an estimate for its cross section and secondly, the statistical significance of the excess can easily be estimated using a bootstrapping-based likelihood-ratio test.

Primary authors: Mr MALMI, Eric (Aalto University); Mr KUUSELA, Mikael (Helsinki Institute of Physics (HIP)); Mr VATANEN, Tommi (Aalto University)

Co-authors: Mr ORAVA, Risto (Helsinki Institute of Physics (HIP)); Mr RAIKO, Tapani (Aalto University)

Presenter: Mr KUUSELA, Mikael (Helsinki Institute of Physics (HIP))

Session Classification: Tuesday 06th - Data Analysis –Algorithms and Tools

Track Classification: Track 2 : Data Analysis - Algorithms and Tools