Geophysical Research Abstracts, Vol. 9, 04580, 2007
SRef-ID: 1607-7962/gra/EGU2007-A-04580


# Machine learning for compliance verification of the Comprehensive Nuclear-Test-Ban Treaty.

J.G. Li(1),**T.J.Stocki**.(2,3),N. Japkowicz.(1), R.K.Ungar(2)

(1) School for Information Technology and Engineering, University of Ottawa, Ottawa, Ontario, Canada, (2) Radiation Protection Bureau, Health Canada, Ottawa, Ontario, Canada.(3) Corresponding Author: Trevor_stocki@hc-sc.gc.ca Fax: (613) 957-1089 Phone:(613) 941-5175

**Abstract** A possible method of weapon detection for the Comprehensive nuclear-Test-Ban-Treaty consists of monitoring the amount of radioxenon in the atmosphere by measuring and sampling the activity concentration of Xe-131m, Xe-133, Xe-133m, and Xe-135 by radionuclide monitoring. Several samples were obtained under different circumstances of nuclear weapon detonation, and are used as training datasets to establish an optimal classification model employing state-of-the-art technologies in machine learning (ML). We conducted a preliminary study involving ML algorithms including Naïve Bayes, Neural Networks, Decision Trees, k-Nearest Neighbours, and Support Vector Machines, that revealed that any noise, uncorrelated features, and interactions in extracted weapon signals will cause difficulties for induction algorithms. We developed a novel feature selection approach that addresses these issues. The method is based on the Gram-Schmidt orthogonalization procedure, and can be used to rebuild the whole feature space such that the resulting features are orthonormal to each other (they do not interact with each other, and each resulting feature is sufficiently correlated with the target. This approach was shown to boost performance in 16 out of 36 experiments where no feature selection was applied (in four cases, by more than 10%), to change nothing in 11 cases and to degrade performance in 9 cases (in only three of these cases, more than 2% degradation occurred, but never more than 4.2%). This method was also shown to obtain an improvement of 4.59% in accuracy over 10 state-of-the-art feature selection methods and no feature selection, on our most challenging data set.

Keywords: Machine Learning, Comprehensive Nuclear-Test-Ban Treaty, radioxenon

monitoring.