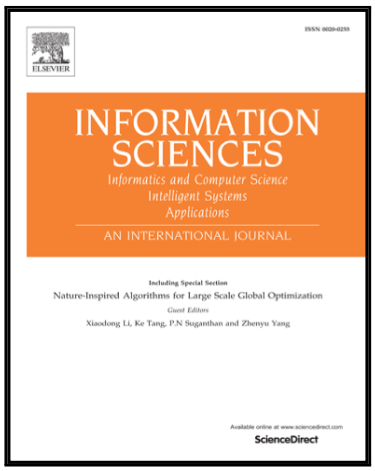

Semi-Supervised Active Learning for Support Vector Machines: A Novel Approach that Exploits Structure Information in Data

Adrian Calma, Tobias Reitmaier, Bernhard Sick

PII: S0020-0255(18)30330-X
DOI: 10.1016/j.ins.2018.04.063
Reference: INS 13606

To appear in: *Information Sciences*

Received date: 27 February 2016
Revised date: 20 April 2018
Accepted date: 21 April 2018

Please cite this article as: Adrian Calma, Tobias Reitmaier, Bernhard Sick, Semi-Supervised Active Learning for Support Vector Machines: A Novel Approach that Exploits Structure Information in Data, *Information Sciences* (2018), doi: 10.1016/j.ins.2018.04.063

Highlights

- Our new approach starts with an empty initial training set.
- Data structure is captured with help of generative, probabilistic mixture models, that are initially trained with variational Bayesian inference in an unsupervised fashion.
- The underlying data models are adapted (revised) during the active learning process with help of class information that becomes available.
- Structure information is considered (1) for the active sample selection by a self-adaptive, multi-criteria selection strategy and (2) for support vector machine training (finding the support vectors) with help of a datadependent kernel.
- Bringing all together results in a practical, effective, and efficient approach that combines active learning and semi-supervised learning for support vector machines.
- Results are shown for 21 publicly available benchmark data sets, including the large MNIST data set.