ADVANCES in UNIVERSUM LEARNING

Vladimir Cherkassky [1] and Sauptik Dhar [2]
[1]Dept. of Electrical & Computer Eng., University of Minnesota, Minneapolis, MN 55455
tel 612 625-9597 fax 612 625-4583 email cherk001@umn.edu
http://www.ece.umn.edu/users/cherkass/predictive_learning/

[2] Research and Technology Center, Robert Bosch LLC, Palo Alto, CA 94304
tel 650 320-1540 email sauptik.dhar@us.bosch.com

SUMMARY:
Most learning methods developed in statistics, machine learning, and pattern recognition assume a standard inductive learning formulation, in which the goal is to estimate a predictive model from finite training data. While this *inductive setting* is very general, there are several emerging non-standard learning settings that are particularly attractive for data-analytic modeling with sparse high-dimensional data. Such recent non-standard learning approaches include transduction, learning using privileged information, universum learning and multi-task learning. This tutorial describes the methodology called Universum learning or learning through contradiction (Vapnik 1998, 2006, Weston et al 2006, Sinz et al 2008). It provides a formal mechanism for incorporating a priori knowledge for binary classification problems. This knowledge is provided in the form of unlabeled Universum data samples, in addition to labeled training samples (under standard inductive setting). The Universum samples belong to the same application domain as training data. However, they do not belong to either class, so they are treated as *contradictions* under a modified SVM-like Universum formulation. Several recent analytical and empirical studies provide ample evidence that Universum learning can improve generalization performance, especially for very ill-posed sparse settings.
This tutorial will present an overview of Universum learning for *binary classification* along with practical conditions for evaluating the effectiveness of Universum learning, relative to standard SVM classifiers (Cherkassky et al, 2011; Cherkassky, 2013). We also describe an extension of Universum SVM to cost-sensitive classification settings (Dhar and Cherkassky, 2012).
The Universum learning methodology is known only for classification setting. It is not clear how to extend the idea of learning through contradiction to other types of learning problems because the notion of 'contradiction' has been originally introduced for binary classification (Vapnik 1998, 2006). In the second part of this tutorial we present general methodology for incorporating Universum into other types of learning problems. For these problems, one can also expect to achieve improved generalization performance by including additional data samples reflecting a priori knowledge about an application domain. In particular, we present new SVM-based formulation for *regression* and *single-class learning* problems that incorporate additional Universum data. Then we briefly discuss computational implementations of these new Universum optimization formulations. We also present several application examples to illustrate advantages of these new Universum formulations, relative to standard SVM regression and single-class SVM. Further, we discuss how the Universum single-class learning can be used for difficult classification problems in changing (nonstationary) environments.

TUTORIAL LENGTH: 2 hours.

INTENDED AUDIENCE: Researchers and practitioners interested in understanding advanced SVM-based methods and applications. Participants are expected to have background knowledge of standard Support Vector Machine (SVM) classifiers.

References

Bai, X. and V. Cherkassky (2008), Gender classification of human faces using inference through contradictions, *in Proc. IJCNN*, Hong Kong, pp. 746–750, 2008

Cherkassky, V. and F. Mulier, Learning from Data, *second edition*, Wiley, 2007

Cherkassky, V., Predictive Learning, http://vctextbook.com/ 2013

Cherkassky, V., Dhar, S., and W. Dai, Practical Conditions for Effectiveness of the Universum Learning, *IEEE Transactions on Neural Networks*,vol.22, no. 8, 1241-1255, 2011.

Dhar, S. and V. Cherkassky, Cost-Sensitive Universum-SVM, Proc. *ICMLA*, 2012

Vapnik, V., Statistical Learning Theory, Wiley, 1998

Vapnik, V., Empirical Inference Science: Afterword of 2006, Springer 2006

Weston, J, Collobert, R., Sinz, F., Bottou, L. and V. Vapnik (2006), Inference with the Universum, *Proc. ICML*.

Sinz, F., O. Chapelle, A. Agarwal, and B. Schölkopf, An analysis of inference with the Universum, *Proc. NIPS-21*, pp. 1–8, 2008.