

Processing challenging data for machine learning

Ireneusz Czarnowski¹

¹*Faculty of Computer Science, Gdynia Maritime University, Poland*

Traditional techniques used for machine learning are not well fit to data, which are difficult, i.e. have a high dimensionality, imbalanced class distribution, are uncompleted, damaged, massive or are streaming in nature. Thus, searching for new and better techniques, methods or approaches suitable for working with such data is a really actual for the machine learning community. In my talk I will present selected research results on working with difficult data. Proprietary approaches for learning from big dataset, as well as, from imbalanced data and data streams will be presented. It will be also presented research results on methods for multi-label learning in the selected telecommunication system, where a problem of data damaged is observed.