Weak Signals: Machine Learning Meets Extreme Value Theory

Stephan Clémençon[1]

[1] <Affiliation 1> Telecom Paris, LTCI, Institut Polytechnique de Paris, France

Abstract:
Rare and extreme events have a major impact on risk management in a wide variety of domains from environmental sciences (heat waves, flooding) to finance and insurance (financial crashes, reinsurance). Furthermore, motivated by a wide variety of applications including fraud detection, the monitoring of complex networks and aviation safety management, unsupervised anomaly detection has recently received much attention in the machine learning community. This important area of machine learning is naturally related to extreme events analysis. AI is expected to provide systems of predictive maintenance for complex infrastructures, such as electricity grids or aircrafts, facilitating the early detection of "weak" signals that announce breakdowns, and serving to plan the replacement of components before their probable failure. As an example, when a complex system is monitored by several physical variables, controlling the false alarm rate is a major issue which can be addressed in the statistical framework of extreme value theory. In the Big Data era, though occurring with very small probability, extreme events are becoming observable. Hence, very large data sets provide the opportunity to develop novel machine learning methods tailored to the predictive analysis of events located far from the 'center' of the distribution. It is the goal of this talk to explain how the framework of multivariate regular variation offers the opportunity to design statistical learning methods in extreme regions. Whereas the majority of machine learning methods aim at exploiting data close to the mean in order to minimize risk functionals like the standard classification error, the mean squared error in regression or the within cluster point scatter in unsupervised learning, algorithms tailored to capture discriminating patterns in the tail behavior, relying on the concept of multivariate regular variation, can be successfully designed. These patterns are generally described by the angular measure, a marginal of the limit distribution of the heavy-tailed random vector of interest that fully captures the dependence structure among multivariate extremes. This talk proposes a survey of several statistical learning methods based on this observation. 1) A dimensionality reduction technique that identifies groups of variables that can be simultaneously large together, while the other variables remain small, providing thus a generally sparse representation of the extremal behavior of a random vector of high dimensionality. 2) A clustering technique for extreme data, as well as a dedicated visualisation method facilitating interpretation. 3) A general approach to learn classifiers that are accurate in extreme regions. Grounds for the validity of these methods will also be provided in the form of generalization bounds.

Keywords:
Multivariate regular variation; extreme value theory; statistical learning; generalization ability; anomaly detection, classification, dimensionality reduction, concentration bounds; empirical angular measure

References:
[1] A Multivariate Extreme Value Theory Approach to Anomaly Clustering and Visualization. With M. Chiapino, V. Feuillard and A. Sabourin. In **Computational Statistics**, 2019.

[2] On Binary Classification in Extreme Regions. With H. Jalalzai and A. Sabourin. In the **Proceedings of NIPS**, 2018.
[3] Anomaly Detection in Extreme Regions via Empirical MV-sets on the Sphere. With A. Thomas, A. Sabourin and A. Gramfort. In the **Proceedings of AISTATS**, 2017.
[4] Sparse Representation of Multivariate Extremes with Applications to Anomaly Detection. With N. Goix and A. Sabourin. In **Journal of Multivariate Analysis**, 2017.
[5] Sparse Representation of Multivariate Extremes with Applications to Anomaly Ranking. With N. Goix & A. Sabourin. In the **Proceedings of AISTATS 2016.**
[6] Learning the dependence structure of rare events: a non asymptotic study. With N. Goix and A. Sabourin. In the **Proceedings of the 2015 COLT Conference**.
[7] Concentration bounds for the empirical angular measure with statistical learning applications. With Hamid Jalalzai, Anne Sabourin, Johan Segers. 2021, **https://arxiv.org/abs/2104.03966**