Rank-based support vector machines for highly imbalanced data using nominated samples

Mohammad Jafari Jozani

(University of Manitoba, Winnipeg, Canada)

21.06.2023, 16.00 (c.t.)

Department of Statistics, Ludwigstr. 33, Room 144
and online via Zoom (Link)
(Meeting-ID: 913-2473-4411; Password: StatsCol22)

We propose a novel approach to address the issue of highly imbalanced binary classification problems using rank information. Our approach utilizes a maxima nominated sampling technique that biases the training sample towards the minority class by using observations with the highest chance of being from the minority class in a small sample of randomly selected units from the underlying population. This sampling technique is based on expert opinion, which has received minimal attention in the machine learning community so far. To incorporate the extra rank information of maxima nominated samples (MaxNS) into the learning process, we introduce novel rank-based Hinge and Logistic loss functions that account for the extra rank information in MaxNS training data sets. We develop MaxNS Support Vector Machines and provide efficient algorithms for solving the proposed learning problems. Numerical studies are performed to validate the efficacy of the methods presented.

Biography:
Mohammad Jafari Jozani is a Professor of Statistics with the Department of Statistics at the University of Manitoba. His research revolves around data-centric approaches for developing fundamental theories, new methodologies and computational tools to solve problems of relevance in a variety of application domains.