



Contribution ID: 69

Type: **not specified**

High-dimensional copula-based classification using truncation and thresholding

Tuesday, 18 May 2021 10:00 (20 minutes)

Bayes classifiers rest on maximising the joint conditional PDF of the feature vector under the class value. The usage of copulae is the most flexible way of fitting joint distributions to data. In recent years, the problem of applying copulae to high dimensions has been approached with Vine copulae. Nevertheless, the application to very high dimensions in the order of several thousands have not yet been studied on large scale in the literature. The present work investigates the feasibility of Bayes classification based on copula modelling in problems with up to 5000 feature components, with a relatively small sample to dimension ratio. To fit Vine copulae successfully in useful computational time in this environment, we use truncation and thresholding. In particular, the potential of thresholding has not yet been studied in classification approaches. We develop approaches for choosing the relevant thresholding levels. Simulation experiments show that the resulting classifiers are strongly competitive with other classifiers as SVM.

Primary authors: Mr WACHTER, Max-Carl (University of Wuerzburg); Mr EASTON, Andrew (University of Wuerzburg); GÖB, Rainer (University of Wuerzburg)

Presenter: Mr WACHTER, Max-Carl (University of Wuerzburg)

Session Classification: Process modelling

Track Classification: Data Science in Process Industries