

Feature evaluation and selection: beyond classification

Marko Robnik-Šikonja¹

¹Faculty of Computer and Information Science, University of Ljubljana

In data science, feature selection algorithms aim to reduce the dimensionality of the data and increase the performance of prediction models. Several successful feature evaluation approaches exist, mostly focusing on classification. We will present ideas of classical supervised approaches (filter, wrapper, and embedded methods), both heuristic and optimization based. We will focus our presentation on recent results in extensions to supervised learning (multi-task, multi-view, and multilabel), as well as unsupervised and semi-supervised learning. We will present important issues in feature subset selection such as stability, redundancy, and higher order interactions.

References

- [1] J. Li, K. Cheng, S. Wang, F. Morstatter, R. P. Trevino, J. Tang and H. Liu, Feature selection: A data perspective, *ACM Computing Surveys*, **50**(6) (2017), p. 94.
- [2] R. Sheikhpour, M. A. Sarram, S. Gharaghani and M. A. Z. Chahooki, A survey on semi-supervised feature selection methods, *Pattern Recognition*, **64** (2017), 141–158.
- [3] J. Tang, S. Alelyani and H. Liu, *Feature Selection for Classification: A review*, *Data Classification: Algorithms and Applications*, CRC Press, London, New York, 2014, p. 37.