

Vencálek, Ondřej; Demni, Houyem; Messaoud, Amor; Porzio, Giovanni C.

On the optimality of the max-depth and max-rank classifiers for spherical data. (English)

Zbl 07217114

Appl. Math., Praha 65, No. 3, 331-342 (2020)

Summary: The main goal of supervised learning is to construct a function from labeled training data which assigns arbitrary new data points to one of the labels. Classification tasks may be solved by using some measures of data point centrality with respect to the labeled groups considered. Such a measure of centrality is called data depth. In this paper, we investigate conditions under which depth-based classifiers for directional data are optimal. We show that such classifiers are equivalent to the Bayes (optimal) classifier when the considered distributions are rotationally symmetric, unimodal, differ only in location and have equal priors. The necessity of such assumptions is also discussed.

MSC:

62H30 Classification and discrimination; cluster analysis (statistical aspects)

62G30 Order statistics; empirical distribution functions

Keywords:

depth-based classifier; von Mises-Fisher distribution; directional data; cosine depth

Software:

ISLR

Full Text: [DOI](#)

References:

- [1] Agostinelli, C.; Romanazzi, M., Nonparametric analysis of directional data based on data depth, *Environ. Ecol. Stat.* 20 (2013), 253-270· doi:10.1007/s10651-012-0218-z
- [2] Batschelet, E., *Circular Statistics in Biology*, Mathematics in Biology. Academic Press, London (1981)
- [3] Bowers, J. A.; Morton, I. D.; Mould, G. I., Directional statistics of the wind and waves, *Appl. Ocean Research* 22 (2000), 13-30· doi:10.1016/S0141-1187(99)00025-5
- [4] Chang, T., Spherical regression and the statistics of tectonic plate reconstructions, *Int. Stat. Rev.* 61 (1993), 299-316· doi:10.2307/1403630
- [5] Demni, H.; Messaoud, A.; Porzio, G. C., The cosine depth distribution classifier for directional data, *Applications in Statistical Computing Studies in Classification, Data Analysis, and Knowledge Organization*. Springer, Cham (2019), 49-60· doi:10.1007/978-3-030-25147-5_4
- [6] Fisher, N. I., Smoothing a sample of circular data, *J. Struct. Geol.* 11 (1989), 775-778· doi:10.1016/0191-8141(89)90012-6
- [7] Ghosh, A. K.; Chaudhuri, P., On maximum depth and related classifiers, *Scand. J. Stat.* 32 (2005), 327-350· doi:10.1111/j.1467-9469.2005.00423.x
- [8] Hubert, M.; Rousseeuw, P.; Segaeert, P., Multivariate and functional classification using depth and distance, *Adv. Data Anal. Classif., ADAC* 11 (2017), 445-466· doi:10.1007/s11634-016-0269-3
- [9] James, G.; Witten, D.; Hastie, T.; Tibshirani, R., *An Introduction to Statistical Learning: With applications in R*, Springer Texts in Statistics 103. Springer, New York (2013)· doi:10.1007/978-1-4614-7138-7
- [10] Kirschstein, T.; Liebscher, S.; Pandolfo, G.; Porzio, G. C.; Ragozini, G., On finite-sample robustness of directional location estimators, *Comput. Stat. Data Anal.* 133 (2019), 53-75· doi:10.1016/j.csda.2018.08.028
- [11] Klecha, T.; Kosiorowski, D.; Mielczarek, D.; Rydlewski, J. P., New proposals of a stress measure in a capital and its robust estimator, Available at <https://arxiv.org/abs/1802.03756> (2018), 24 pages
- [12] Kosiorowski, D., About phase transitions in Kendall's shape space, *Acta Univ. Lodz., Folia Oeconomica* 206 (2007), 137-155
- [13] Leong, P.; Carlile, S., Methods for spherical data analysis and visualization, *J. Neurosci. Methods* 80 (1998), 191-200· doi:10.1016/S0165-0270(97)00201-X
- [14] Ley, C.; Sabbah, C.; Verdebout, T., A new concept of quantiles for directional data and the angular Mahalanobis depth, *Electron. J. Stat.* 8 (2014), 795-816· doi:10.1214/14-EJS904
- [15] Liu, R. Y., On a notion of data depth based on random simplices, *Ann. Stat.* 18 (1990), 405-414· doi:10.1214/aos/1176347507

- [16] Liu, R. Y.; Singh, K., Ordering directional data: Concepts of data depth on circles and spheres, *Ann. Stat.* 20 (1992), 1468-1484· doi:10.1214/aos/1176348779
- [17] Makinde, O. S.; Fasoranbaku, O. A., On maximum depth classifiers: Depth distribution approach, *J. Appl. Stat.* 45 (2018), 1106-1117· doi:10.1080/02664763.2017.1342783
- [18] Mardia, K. V.; Jupp, P. E., *Directional Statistics*, Wiley Series in Probability and Statistics. John Wiley & Sons, Chichester (2000)· doi:10.1002/9780470316979
- [19] Paindaveine, D.; Verdebout, T., Optimal rank-based tests for the location parameter of a rotationally symmetric distribution on the hypersphere, *Mathematical Statistics and Limit Theorems Springer, Cham* (2015), 249-269· doi:10.1007/978-3-319-12442-1_4
- [20] Pandolfo, G.; D'Ambrosio, A.; Porzio, G. C., A note on depth-based classification of circular data, *Electron. J. Appl. Stat. Anal.* 11 (2018), 447-462· doi:10.1285/i20705948v11n2p447
- [21] Pandolfo, G.; Paindaveine, D.; Porzio, G. C., Distance-based depths for directional data, *Can. J. Stat.* 46 (2018), 593-609· doi:10.1002/cjs.11479
- [22] Saw, J. G., A family of distributions on the (m) -sphere and some hypothesis tests, *Biometrika* 65 (1978), 69-73· doi:10.1093/biomet/65.1.69
- [23] Small, C. G., Measures of centrality for multivariate and directional distributions, *Can. J. Stat.* 15 (1987), 31-39· doi:10.2307/3314859
- [24] Tukey, J. W., Mathematics and the picturing of data, *Proceedings of the International Congress of Mathematicians Canad. Math. Congress, Montreal* (1975), 523-531
- [25] Vencálek, O., Depth-based classification for multivariate data, *Austrian J. Stat.* 46 (2017), 117-128· doi:10.17713/ajs.v46i3-4.677

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.