Explaining probabilistic prediction over the simplex with Shapley compositions

Anonymous Authors[1]

Abstract

The concept of Shapley value has been widely use for measuring the contribution of each features on a machine learning model's prediction. However, this has been designed for one-dimensional function's codomain. For multiclass probabilistic classifier, where the output is a discrete probability distribution over the set of more than two possible classes, the output domain is a multidimensional simplex. In this case, people usually apply the concept of Shapley value on each each output dimension one-by-one, in an implicit one-vs-rest setting ignoring the compositional nature of the output distribution. Using the Aitchison geometry of the simplex, coming from the field of compositional data analysis, this paper present a multidimensional extention of the concept of Shapley value, names Shapley composition, for explaining probabilistic predictions one the simplex in machine learning.

1. Introduction

References

Author, N. N. Suppressed for anonymity, 2021.

Duda, R. O., Hart, P. E., and Stork, D. G. Pattern Classification. John Wiley and Sons, 2nd edition, 2000.

Kearns, M. J. Computational Complexity of Machine Learning. PhD thesis, Department of Computer Science, Harvard University, 1989.

Langley, P. Crafting papers on machine learning. In Langley, P. (ed.), Proceedings of the 17th International Conference on Machine Learning (ICML 2000), pp. 1207–1216, Stanford, CA, 2000. Morgan Kaufmann.

Michalski, R. S., Carbonell, J. G., and Mitchell, T. M. (eds.). Machine Learning: An Artificial Intelligence Approach, Vol. I. Tioga, Palo Alto, CA, 1983.

Mitchell, T. M. The need for biases in learning generalizations. Technical report, Computer Science Department, Rutgers University, New Brunswick, MA, 1980.

Newell, A. and Rosenbloom, P. S. Mechanisms of skill acquisition and the law of practice. In Anderson, J. R. (ed.), Cognitive Skills and Their Acquisition, chapter 1, pp. 1–51. Lawrence Erlbaum Associates, Inc., Hillsdale, NJ, 1981.

Samuel, A. L. Some studies in machine learning using the game of checkers. IBM Journal of Research and Development, 3(3):211–229, 1959.

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.

A. You can have an appendix here.

You can have as much text here as you want. The main body must be at most 8 pages long. For the final version, one more page can be added. If you want, you can use an appendix like this one.

The `\onecolumn` command above can be kept in place if you prefer a one-column appendix, or can be removed if you prefer a two-column appendix. Apart from this possible change, the style (font size, spacing, margins, page numbering, etc.) should be kept the same as the main body.