

Set-valued classification under fairness constraints

Keywords : Supervised learning, Algorithmic fairness, Set-valued classification

Context : Multiclass classification is one of the most studied statistical framework, arising in many fields which range from medical applications to social studies. In the last decade, several challenges have emerged in multi-class classification. This internship aims to make a particular focus on two of them.

On one hand, due to high ambiguity between classes inherent in large scale multi-class datasets, it is more relevant to consider set-valued classifiers rather than usual single-output classifiers [1, 4, 5].

On the other hand, mitigating bias in data is an active research field in the machine learning community. This is premised on the fact that learning algorithms may inherit bias in the data during the training process, leading to undesired knock-on effect on future decisions. In particular, severe conflicts may arise with ethical criteria of the modern society using algorithms that only have a purpose of prediction accuracy. Algorithmic fairness, which has been emerging in the last few years, try to give a solution to the problem of mitigating the bias in data [2, 6, 7].

The aim of this internship is to study set-valued classification under fairness constraint.

Objectives : A preliminary work of [3] considers multi-class classifiers under the popular but limited demographic parity constraint. One of the biggest challenges of this internship is to extend the methodology of [3] to set-valued classification framework as well as to consider other fairness measures. In particular, this generalization relies on statistical learning theory and non differentiable optimization techniques. Ultimately, the objectives of this internship are to

1. build algorithms dedicated to set-valued classification in the context of fairness ;
2. implement the developed procedure in `Python` language ;
3. establish its theoretical and numerical properties.

Supervisors : Christophe Denis (LPSM, Sorbonne-Université), Mohamed Hebiri (LAMA, Université Gustave Eiffel).

Required skills : M2 level trainee in statistics/machine learning/optimization. Python programming. Applicants should send a CV and transcripts of the last two years to cdenis@lpsm.paris and mohamed.hebiri@univ-eiffel.fr

Practical information :

- April-September 2024
- Location : LPSM, Sorbonne Université
- Grant from SCAI

References

- [1] Barber R., Candes E., Ramdas A., and Tibshirani, R.J The limits of distribution-free conditional predictive inference. Information and Inference : A Journal of the IMA, vol. 10(2), pp. 455-482 (2021)
- [2] Calders T., Kamiran F., and Pechenizkiy M. Building classifiers with independency constraints. IEEE international conference on Data mining (2009)
- [3] Denis C, Elie, R., Hebiri, M., and Hue, F. Multiclass classification under fairness guarantees. Preprint (2022)

- [4] Denis C and Hebiri M. Confidence sets with expected sizes for multiclass classification. Journal of Machine Learning Research, vol. 18(102), pp. 1-28 (2017)
- [5] Vovk V., Gammerman A., and Shafer G. Algorithmic learning in a random world. Springer Science & Business Media (2005)
- [6] Zafar M., Valera I., Gomez-Rodriguez M., and Gummadi, K. Fairness Constraints : A Flexible Approach for Fair Classification. Journal of Machine Learning Research, vol. 20(75), pp. 1-42 (2019)
- [7] Zemel R., Wu Y., Swersky K., Pitassi T., and Dwork C. Learning fair representations. International Conference on Machine Learning (2013)