**Doctoral School of Information and Biomedical Technologies**
**Polish Academy of Sciences (TIB PAN)**

---

## SUBJECT:

Detecting and modelling test data distribution shift in supervised learning

## SUPERVISOR:

Dr. Paweł Teisseyre, teisseyrep@ipipan.waw.pl, Institute of Computer Science, Polish Academy of Sciences, 01-248 Jana Kazimierza 5, Warsaw, Poland

## DESCRIPTION:

Traditional supervised learning methods are based on the classic assumption that training and test data are drawn from the same probability distribution. However, in many real applications the above assumption is not met, i.e. the model is trained on data from a certain distribution (source distribution), and then applied to new data that may come from another distribution (target distribution). For example, in medical applications, a model predicting the occurrence of a disease can be trained on data about patients from one country, and then we use it to predict the disease for patients from another country. In object recognition in images, the training set may contain objects captured in a different scenery than the objects in the test data.

There are several types of distribution shift [1]. Label shift [2] involves changing the prior distribution of a class variable or multiple class variables, while the conditional distribution of features given the class variable does not change. Covariate shift [3] involves changing the marginal distribution of features, while the posterior distribution of the class variable does not change. We can also consider the most general situation in which both the marginal and conditional distributions change [4]. A closely related task is to verify whether a single observation in the test set comes from a distribution other than the training distribution, the so-called out-of-distribution detection [5-6].

The proposed topic includes analyzing methods for detecting distribution shifts, including out-of-distribution detection as well as model correction methods that take into account the change in distribution [1-4 ]. We plan to apply the developed methods in various fields, in particular in the problem of recognizing objects in images.

Candidate should have M.Sc. in Mathematics, Computer Science or Engineering, have experience in Machine Learning and Statistics, including both its mathematical and computational aspects, and possess sufficient computing skills (knowledge of scikit-learn and pytorch libraries is obligatory) to effectively implement proposed methods. Scientific curiosity and eagerness to learn are essential.

Candidate should contact the author of the proposal before formal submission of documents (teisseyrep@ipipan.waw.pl).

## BIBLIOGRAPHY:


[1] B. Zadrozny, Learning and evaluating classifiers under sample selection bias, Proceedings of the International Conference on Machine Learning, ICML, 2004.

[2] Z. C. Lipton, Y. Wang, A. J. Smola, Detecting and Correcting for Label Shift with Black Box Predictors, Proceedings of the International Conference on Machine Learning, ICML, 2014.

[3] S. Bickel, M. Bruckner, T. Scheffer, Discriminative Learning Under Covariate Shift, Journal of Machine Learning Research, 2009

[4] S. Maity, M. Yurochkin and M. Banerjee and Y. Sun, Understanding new tasks through the lens of training data via exponential tilting, Proceedings of the International Conference on Learning Representations, ICRL, 2023.

[5] J. Yang, K. Zhou, Y. Li, Z. Liu, Generalized Out-of-Distribution Detection: A Survey, https://arxiv.org/abs/2110.11334.

[6] W. Liu, X. Wang, J. Ownes, Y. Li, Energy-based Out-of-distribution Detection, Advances in Neural Information Processing Systems, NIPS, 2020.