Title: **Hyper-parameter Tuning of Machine Learning Algorithms**

Supervisor: Tomáš Horváth (`horvathtomi1976@gmail.com`)

Affiliation: Data Science Department, Faculty of Informatics, ELTE

Hyper-parameter (HP) tuning of machine learning (ML) algorithms is, in general, treated as a black-box optimization problem [1] which objective function[1] $f : \mathcal{A} \times \mathcal{D} \times \mathcal{H} \rightarrow \mathbb{R}$ captures the predictive performance of the algorithm $a \in \mathcal{A}$ with the HP setting $\mathbf{h} = (h_1, h_2, \dots, h_k) \in \mathcal{H}$ on the dataset $\mathbf{D} \in \mathcal{D}$ where $\mathcal{A}$ is the set of ML algorithms, $\mathcal{D}$ is the set of datasets and $\mathcal{H} = \mathcal{H}_1 \times \mathcal{H}_2 \times \cdots \times \mathcal{H}_k$ is the space of admissible values (usually defined by some constraints) of HPs for the algorithm a. The task of HP tuning is, given a, $\mathcal{H}$ and $\mathbf{D}$, to find $\mathbf{h}^\star \in \mathcal{H}$ such that

$$\mathbf{h}^\star = \underset{\mathbf{h} \in \mathcal{H}}{arg\,max}\, f(a, \mathbf{D}, \mathbf{h}) \tag{1}$$

Various approaches to HP tuning were developed ranging from simpler grid [2] or random search [3], direct search methods [4] through more sophisticated approaches such as Evolutionary Algorithms [5], Particle Swarm Optimization [6], Sequential Model-based Optimization [7], Bayesian methods [8], etc.

The goal of the work is to create a sandbox framework for test and comparison of HP tuning approaches including the following steps:

- Search for freely available tools and their installation
- Implement not freely available approaches
- Download real-world and generate sythetic benchmark datasets
- Propose a sound experimental methodology for testing and comparison
- Evaluate and analyze the obtained results

References

[1] D. R. Jones, M. Schonlau, and W. J. Welch (1998) *Efficient global optimization of expensive black-box functions.* Journal of Global Optimization, vol. 13, no. 4, p:455–492.

[2] I. Braga, L. P. do Carmo, C. C. Benatti, and M. C. Monard (2013). *A note on parameter selection for support vector machines.* In Lecture Notes in Computer Science, vol. 8266, p:233–244.

[3] J. Bergstra and Y. Bengio (2012). *Random search for hyper-parameter optimization.* Journal of Machine Learning Research, vol. 13, pp. 281–305.

[4] F. Hutter, H. Hoos, K. Leyton-Brown, and T. Stützle (2009). *Paramils: an automatic algorithm configuration framework.* Journal of Artificial Intelligence Research, no. 36, p:267–306.

[1] Any, even multi-objective, performance measure can be utilized for f.

[5] F. Friedrichs and C. Igel (2005). *Evolutionary tuning of multiple svm parameters.* Neurocomputing, vol. 64, p:107–117.

[6] A. Kalos (2005). *Automated neural network structure determination via discrete particle swarm optimization (for non-linear time series models).* In 5th WSEAS International Conference on Simulation, Modeling and Optimization, p:325–331.

[7] F. Hutter, H. H. Hoos, and K. Leyton-Brown (2011). *Sequential model-based optimization for general algorithm configuration.* In Lecture Notes in Computer Science, vol. 6683, Springer, p:507–523.

[8] C. Thornton, F. Hutter, H. H. Hoos, and K. Leyton-Brown (2013). *Auto-WEKA: Combined selection and hyperparameter optimization of classification algorithms.* Proceedings of the International Conference on Knowledge Discovery in Databases 2013, pp:847–855.