

# Strojové učení jako špatně podmíněná úloha

## Machine Learning as Ill-Posed Task

Školitel: doc. Ing. Jaromír Kukal, Ph.D.

V oblasti strojového učení se běžně setkáváme s daty, u kterých je počet příznaků výrazně vyšší než počet vzorů k učení. Navíc další příznaky mohou vzniknout různými nelineárními transformacemi dat (skrytá vrstva ANN, systematický preprocesing, jádrové transformace). Následné určení vah jednotlivých příznaků pro klasifikaci nebo predikci je špatně podmíněnou úlohou ve statistickém slova smyslu. Klasickým řešením je regularizace úlohy s využitím norem  $L_\infty$  (robustní klasifikace a regrese),  $L_2$  (max margin a ridge regrese),  $L_1$  (prořezávání a lasso regrese) a  $L_0$  (minimalizace počtu příznaků).

Cílem práce je vyvinout alternativy k  $L_0$  regularizaci, které by optimalizovaly výběr příznaků s využitím statistických modelů, metody maximální věrohodnosti a informačních kritérií. Uvedeného cíle lze dosáhnout kombinací vhodně zvolených modelů, stochastických metod učení a binárních optimalizačních heuristik. Doktorand by měl mít znalosti matematické analýzy, matematické statistiky, optimalizačních heuristik a programování (Matlab, C, Python).

- Vidaurre, D., Bielza, C., & Larranaga, P. (2013). A survey of L1 regression. *International Statistical Review*, 81(3), 361-387.
- Huang, J., Jiao, Y., Liu, Y., & Lu, X. (2018). A constructive approach to L0 penalized regression. *The Journal of Machine Learning Research*, 19(1), 403-439.
- Klimt, M., Kukal, J., & Mojzeš, M. (2013). Lévy flights in binary optimization. *Archives of Control Sciences*, 23(4).
- Kukal, J., & Mojzeš, M. (2018). Quantile and mean value measures of search process complexity. *Journal of Combinatorial Optimization*, 35, 1261-1285.
- Chawla, M., & Duhan, M. (2018). Levy flights in metaheuristics optimization algorithms—a review. *Applied Artificial Intelligence*, 32(9-10), 802-821.