



# **Integrating elements of data science into high-school teaching: Naïve Bayes-classification algorithm and programming in Python**

ÖDÖN VANC SÓ AND PÉTER PRINCZ

*Abstract.* Probability theory and mathematical statistics are traditionally one of the most difficult chapters of mathematics to teach. One of the authors, Péter Princz has experience in teaching various topics via computer programming of the problem at hand as a class activity. The proposed method is to involve programming as a didactic tool in hard-to-teach topics. The intended goal in this case is to implement a naïve Bayes-classifier algorithm in Python and demonstrate the machine-learning capabilities of it by applying it to a real-world dataset of edible or poisonous mushrooms. The students would implement the algorithm in a playful and interactive way. The proposed incremental development process aligns well with the spirit of Tamás Varga who considered computers as modern tools of experimental problem solving as early as in the 1960s.

*Key words and phrases:* Bayesian classifier, computer programming, guided discovery learning.

*MSC Subject Classification:* 97D40, 97D50, 97K50, 97K99, 97M60, 97P40, 97P50, 97U50.

ÖDÖN VANCÓSÓ

ELTE, FACULTY OF SCIENCE, INSTITUTE OF MATHEMATICS  
1117 BUDAPEST, PÁZMÁNY PÉTER SÉTÁNY 1/C.

*E-mail:* [vancso.odon@gmail.com](mailto:vancso.odon@gmail.com)

PÉTER PRINCZ

ELTE, FACULTY OF SCIENCE, INSTITUTE OF MATHEMATICS  
1117 BUDAPEST, PÁZMÁNY PÉTER SÉTÁNY 1/C

*E-mail:* [princzp@caesar.elte.hu](mailto:princzp@caesar.elte.hu)