

Generalized Naive Bayes with continuous explanatory variables

lundi 28 juillet 2025 18:00 (30 minutes)

Naive Bayes is one of the most widely used machine learning algorithms, appreciated for its simplicity, efficiency, and ease of interpretation—qualities that make it appealing across various fields. However, Naive Bayes operates under the assumption that the explanatory variables are conditionally independent given the class label, an assumption that often does not hold true in practice.

To address this limitation, we propose a method that relaxes the independence assumption by allowing certain dependencies between the explanatory variables. The central idea of our approach is to find a structure that best approximates the joint probability distribution of the data while minimizing the Kullback-Leibler divergence. We introduce a greedy algorithm that gradually builds this structure step by step.

Initially, the algorithm was designed for cases with discrete variables. We are now adapting it for situations where the explanatory variables follow a joint multivariate Gaussian probability distribution. Furthermore, we extend our approach to accommodate more general continuous variables.

Finally, we evaluate the performance of our method against other classification techniques using real-world datasets. Our method is also capable of identifying the most informative set of explanatory variables.

Author: M. PAPP, Ábrahám (Budapest University of Technology and Economics)

Co-auteurs: Dr SZILÁGYI, Botond (Budapest University of Technology and Economics); KOVÁCS, Edith Alice (Budapest University of Technology and Economics)

Orateur: KOVÁCS, Edith Alice (Budapest University of Technology and Economics)

Classification de Session: ML

Classification de thématique: Machine learning