

Application of data mining techniques in animal husbandry

Summary

While summing up the results of the analyzes it should be noted that parallel application of two statistical methods, the method of decision trees and multivariate analysis of variance, have indicated the same factors that determine the lactose content in cow's milk: successive lactation, day of milking testing, herd, the participation of HF breed, calving season, year of birth, CASK protein genotype. Based on the constructed model of the decision tree, it can be concluded that a high lactose content may be expected in the primiparas milk, after 17 days of milking, with the participation of HF at the level of at least 74.5%. However, a subsequent production period of more than 2 lactations, promotes the lower lactose content in milk, especially when milk samples, collected from cows with the participation of HF of at least 91.5%, are obtained from flocks with living and feeding conditions similar to those in the herd K. The graphical models of trees, created in the similar way to those described in the research, may be developed for virtually every feature. Therefore, they may become the tool indicating a good system of levels of genetic and environmental factors and provide the high level of features that are useful for breeders.

KEY WORDS: classification trees, modeling, performance traits