89

Count Regression and Machine Learning Techniques for Zero-Inflated Overdispersed Count Data: Application to Ecological Data

The aim of this study is to investigate overdispersion problem that is rampant in ecological count data. In order to explore this problem, we consider the most commonly used count regression models; the Poisson, negative binomial, zero-inflated Poisson and zero-inflated negative binomial models. The performance of these count regression models is compared with the four proposed machine learning (ML) regression techniques; random forest, support vector machines, k-nearest neighbors and artificial neural networks. The mean absolute error (MAE) was used to compare the performance of count regression models and ML regression models. The results suggest that ML regression models performs better compared to count regression models. The performance shown by ML regression techniques is a motivation for further research in improving methods and applications in ecological studies.