Stacking model as a method for genomic prediction

Sunhee Kim, Dongju Lee, and Chang-Young Lee
The Department of Industrial Engineering, Kongju National University

김선희, 이동주, 이창용
공주대학교 산업공학과

As genomic selection emerges as a promising breeding method for plants and animals, numerous methods have been introduced and applied to various real and simulated data sets. Research suggests that no single method is universally better than others; rather, performance is highly dependent on the characteristics of the data and the nature of the prediction task. This study proposes a stacking model as a computational method for genome prediction. The proposed model belongs to the ensemble methods in machine learning and takes a different approach from conventional methods. It integrates base models to explore the collective knowledge from them and uses the meta-model to achieve better performance. The proposed stacking was applied to open-access rice, maize, barley, mice, and millet resources. We found that the proposed method yielded a lower or comparable mean squared error in predicting phenotypes compared to the current methods. In addition, the proposed method showed greater resistance to overfitting than the current methods. In summary, the proposed stacking model integrates currently available methods to achieve stable and better performance.

Keywords: genomic prediction, stacking model, ensemble methods, mean squared error, overfitting