

DEVELOPMENT OF PREDICTIVE MODEL FOR THE CLASSIFICATION OF GENES ASSOCIATED WITH ABIOTIC STRESS-RESISTANT TRAITS IN RICE USING DEEP LEARNING

ABSTRACT

Rice (*Oryza sativa* L.) is a well-studied model organism for plant research. The recent developments in high throughput omics data have produced a significant amounts of rice omics data. To analyze these data, methods are needed such as classification of genes associated with biotic and abiotic stress-related traits. This study aims to develop deep learning based predictive models for the classification of genes associated with abiotic stress-resistant traits in rice. In this study, the rice dataset from the previous study namely “Kok, 2022” have been utilized. 20 descriptors from the four main features (Expression type, Expression value, Co-expression network and PP1 network) were used to build the predictive models with three deep learning classifiers namely Multilayer Perceptron (MLP), Convolutional Neural Network (CNN) and Radial Basis Function Network (RBFN). The performance of the predictive models was measured using classification accuracy and AUROC. Two types of classification for the predictive models were built which includes the trait classification (classification of associated traits for each gene) and gene classification (classification of associated genes for each trait). MLP showed the highest accuracy in trait classification with 58.6% while CNN showed the highest accuracy in gene classification with 99.1% at 140 training epochs. Feature selection before classifying has brought better performance to gene classification with accuracy 99.2% but not for trait classification with 58.5%. With the development of such models, agriculture sectors and geneticists not only can yield better rice varieties, but it also speeds up the process of screening large-scale rice genomic data.

Keywords: Rice, Abiotic stress-resistant traits, Trait classification, Gene classification, Deep learning