Predicting carboxymethyl cellulase assay (CMCase) production using artificial neural network and explicit feature selection approach

Ullah Miah, M. Saef[a]; Sulaiman, Junaida[a]; Zamli, Kamal Z.[a]; Samiur Rashid, Shah[b]; Khan Chowdhury, Ahmed Jalal[c]

[a] Universiti Malaysia Pahang, Faculty of Computing, Pekan, Malaysia

[b] University of Science and Technology Chittagong (USTC), Department of Biochemistry and Biotechnology, Chittagong, Bangladesh

[c] International Islamic University Malaysia, Kulliyyah of Science, Kuantan, Malaysia

ABSTRACT

This paper presents a method for predicting carboxymethyl cellulase (CMCase) production using artificial neural networks (ANNs) and an explicit feature selection approach. A dataset of CMCase production experiments was collected, and an explicit feature selection approach was applied to select the most relevant features for CMCase production prediction. The ANN model was trained using both the selected features and all available features of the CMCase production data. The results showed that the explicit feature selection approach improved the performance of the ANN model in terms of prediction accuracy compared to using all the features available in the dataset. The main effect analysis (MEA) was found to be the best method for selecting the explicit features for predicting CMCase production. The ANN model trained using the MEA identified features, achieved 96.3% R2 score and a MAE of 0.057 and a MSE of 0.035. The proposed method is an effective approach for predicting CMCase production and can be used to optimize CMCase production and reduce costs in various industries.

KEYWORDS

Artificial neural network; CMCase; Enzyme production prediction; Main effect analysis feature selection; Random forest regressor feature selection

ACKNOWLEDGMENT

The work reported in this paper is funded by TransDisciplinary Research Grant Scheme from the Ministry of Higher Education Malaysia titled: An Artificial Neural Network Sine Cosine Algorithm-based Hybrid Prediction Model for the Production of Cellulose Nanocrystals from Oil Palm Empty Fruit Bunch TRGS/1/2018/UMP/01/1/4 (RDU1918014)