

DeepEC: fine-quality prediction of enzyme commission numbers using deep learning

김예지, 김현욱, 류재용, 이상엽[†]
한국과학기술원
(leesy@kaist.ac.kr[†])

Fine-quality enzyme commission (EC) numbers is essential in order to obtain accurate understanding of enzyme functions. To date, several prediction tools for EC number have been developed, but their performance needs further improvement to efficiently process a tremendous volume of protein sequences. Here, we present a deep learning-based computational framework, DeepEC, that predicts EC numbers for protein sequences precisely in a high-throughput manner. DeepEC implements 3 convolutional neural networks (CNNs) for the EC number prediction, and also conducts homology analysis for EC numbers, which cannot be identified by the CNNs. Comparative analyses against 5 representative EC number prediction tools reveal that DeepEC lead to the most precise prediction, and is the fastest and the lightest regarding to the required disk space. Moreover, DeepEC sensitively detects mutated domains/binding site residues in protein sequences.