Jordan University of Science and Technology

A Framework for Predicting Proteins 3D Structures

Authors: Rehab Duwairi, Amal Kassawneh

Abstract: This paper proposes a framework for predicting protein three dimensional structures from their primary sequences. The proposed method utilizes the natural multi-label and hierarchical intrinsic nature of proteins to build a multi-label and hierarchical classifier for predicting protein folds. The classifier predicts protein folds in two stages, at the first stage, it predicts the protein structural class, and in the second stage, it predicts the protein fold. When comparing our technique with SVM, Na?ve Bayes, and Boosted C4.5 we get a higher accuracy more than SVM and better than Na?ve Bayes when using the composition, secondary structure and hydrophobicity feature attributes, and give higher accuracy than C4.5 when using composition, secondary structure, hydrophobicity, and polarity feature attributes. MuLAM was used as a basic classifier in the hierarchy of the implemented framework. Two major modifications were made to MuLAM, namely: the pheromone update and term selection strategies of MuLAM were altered.