

DISSVM: a novel predictor of intrinsically disordered regions in protein

Bortolami O.(1), Tosatto S.C.E. (1)

(1) Dept. of Biology & CRIBI Biotech Centre, Universita' di Padova

Motivation

The prediction of intrinsically disordered regions in proteins from sequence is increasingly becoming of interest, as the presence of many such regions in the complete genome sequences are discovered. Recent reports also indicate disordered regions to be conserved in evolution and to have a distinct preference for certain functional classes.

Methods

A machine learning approach based on a support vector machine (SVM) representation is developed to discriminate disordered regions from sequence is developed and tested. The SVM is trained and benchmarked on two sets, representing long and short disordered regions. A preliminary version of DISSVM was entered into the recent CASP-6 experiment (URL: <http://predictioncenter.llnl.gov/casp6/>) as part of group Casplta (casp id# 0096).

Results

Benchmarking of DISSVM shows similar performance to the state of the art DISOPRED predictor of D. Jones and co-workers. The results for CASP-6 placed the preliminary version of DISSVM, enhanced by manual intervention, at rank two out of 20 participating groups.

Contact email: silvio@cribi.unipd.it