# ScholarWorks

# Protein Fold Recognition Using Adaboost Learning Strategy

Login (/login)

- IUPUI ScholarWorks Repository
- →
- School of Informatics and Computing
- →
- Informatics Theses and Dissertations
- →
- Informatics Graduate Theses and PhD Dissertations
- →
- View Item

# Protein Fold Recognition Using Adaboost Learning Strategy

Su, Yijing

Name: Su_11-15.pdf
Size: 296.7Kb
Format: PDF

View/Open

| | |
|---|---|
| Permanent Link: | http://hdl.handle.net/1805/2267 |
| Date: | 2010-09-29 |
| Keywords: | Adaboost ; Recognition ; Learning Strategy ; Protein Fold |
| Sponsorship: | Jeffrey Huang, Ph.D.,Chair;Yuni Xia, Ph.D.; Jiang Yu Zheng,Ph.D. |

## Abstract:

Protein structure prediction is one of the most important and difficult problems in computational molecular biology. Unlike sequence-only comparison, protein fold recognition based on machine learning algorithms attempts to detect similarities between protein structures which might not be accompanied with any significant sequence similarity. It takes advantage of the information from structural and physic properties beyond sequence information. In this thesis, we present a novel classifier on protein fold recognition, using AdaBoost algorithm that hybrids to k Nearest Neighbor classifier. The experiment framework consists of two tasks: (i) carry out cross validation within the training dataset, and (ii) test on unseen validation dataset, in which 90% of the proteins have less than 25% sequence identity in training samples. Our result yields 64.7% successful rate in classifying independent validation dataset into 27 types of protein folds. Our experiments on the task of protein folding recognition prove the merit of this approach, as it shows that AdaBoost strategy coupling with weak learning classifiers lead to improved and robust performance of 64.7% accuracy versus 61.2% accuracy in published literatures using identical sample sets, feature representation, and class labels.

## This item appears in the following Collection(s)

- Informatics Graduate Theses and PhD Dissertations (/handle/1805/303)

Show Statistical Information (#)

# My Account

- Login
- Register

# Statistics

- Most Popular Items
- Statistics by Country
- Most Popular Authors

About Us (/page/about) | Contact Us (/contact) | Send Feedback (/feedback)

  (/htmlmap)

## FULFILLING *the* PROMISE