Abstract

As biomedical research and healthcare continue to progress in the genomic/post genomic era, a number of important challenges and opportunities exist in the broad area of bioinformatics. In the broader context, the key challenges to bioinformatics essentially all relate to the current flood of raw data, aggregate information, and evolving knowledge arising from the study of the genome and its manifestation.

Protein structure determination and prediction has been a focal research subject in life sciences due to the importance of protein structure in understanding the biological and chemical activities of organisms. The experimental methods used to determine the structures of proteins demand sophisticated equipment and time. A host of computational
methods are developed to predict the location of secondary structure elements in proteins for complementing or creating insights into experimental results.

The present work focuses on secondary structure prediction of proteins. The data mining model is implemented to predict the various parameters related to the secondary structure. These parameters include the alpha helix, beta sheets and hairpin turn. Cluster analysis is used to implement the secondary structure prediction.

Reference

- Chen Yonghui, Reilly Kevin D., Sprague Alan P., Guan Zhijie,(2006) “SEQOPTICS: a protein sequence clustering system” Symposium of Computations in Bioinformatics and Bioscience (SCBB06) in conjunction with the International Multi-Symposiums on Computer and Computational Sciences 2006 (IMSCCS|06), pp 1-5.
- George Tzanis, Christos Berberidis, and Ioannis Vlahavas (2002) “Biological Data Mining” Department of Informatics, Aristotle University of Thessaloniki, Greece, pp 1-8.

Index Terms

Computer Science  Bioinformatics
Key words

Data mining
Cluster analysis
Protein structure prediction