Protein Structure Prediction and Interpretation with Support Vector Machines and Decision Trees

Yi Pan
Department of Computer Science
Georgia State University, USA

Abstract:

Prediction of protein structures from protein sequences using computers is an important step to discover proteins' 3D conformation structures and their functions and hence has profound theoretical and practical significance in areas such as protein engineering and drug design. In this talk, we will discuss our new results in protein secondary structure and Transmembrane protein prediction using Support Vector Machines. We will also discuss how to use a combination of Support Vector Machine and Decision Tree to understand how a prediction is reached through rule extraction. Clearly, a good interpretation is useful for guiding biological experiments and may lead further prediction improvement. A novel approach of rule clustering for super-rule generation will also be briefly discussed.

Bio:

Dr. Yi Pan is Chair and Full Professor in the Department of Computer Science at Georgia State University, USA. Dr. Pan received his B.Eng. and M.Eng. degrees in computer engineering from Tsinghua University, China, in 1982 and 1984, respectively, and his Ph.D. degree in computer science from the University of Pittsburgh, USA, in 1991. Dr. Pan's research interests include parallel and distributed computing, optical networks, wireless networks, and bioinformatics. Dr. Pan has published more than 80 journal papers with 29 papers published in various IEEE journals. In addition, he has published over 100 papers in refereed conferences (including IPDPS, ICPP, ICDCS, INFOCOM, and GLOBECOM). He has also co-edited 18 books (including proceedings) and contributed several book chapters. He has received many awards from agencies such as NSF, AFOSR, JSPS, IISF and Mellon Foundation. Dr. Pan has served as an editor-in-chief or editorial board member for 15 journals including 5 IEEE Transactions and a guest editor for 7 special issues. He has organized several international conferences and workshops and has also served as a program committee member for several major international conferences such as INFOCOM, GLOBECOM, ICC, IPDPS, and ICPP.