World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:14, No:12, 2020

Estimation of Transition and Emission Probabilities

Authors : Aakansha Gupta, Neha Vadnere, Tapasvi Soni, M. Anbarsi

Abstract : Protein secondary structure prediction is one of the most important goals pursued by bioinformatics and theoretical chemistry; it is highly important in medicine and biotechnology. Some aspects of protein functions and genome analysis can be predicted by secondary structure prediction. This is used to help annotate sequences, classify proteins, identify domains, and recognize functional motifs. In this paper, we represent protein secondary structure as a mathematical model. To extract and predict the protein secondary structure from the primary structure, we require a set of parameters. Any constants appearing in the model are specified by these parameters, which also provide a mechanism for efficient and accurate use of data. To estimate these model parameters there are many algorithms out of which the most popular one is the EM algorithm or called the Expectation Maximization Algorithm. These model parameters are estimated with the use of protein datasets like RS126 by using the Bayesian Probabilistic method (data set being categorical). This paper can then be extended into comparing the efficiency of EM algorithm to the other algorithms for estimating the model parameters, which will in turn lead to an efficient component for the Protein Secondary Structure Prediction. Further this paper provides a scope to use these parameters for predicting secondary structure of proteins using machine learning techniques like neural networks and fuzzy logic. The ultimate objective will be to obtain greater accuracy better than the previously achieved.

Keywords : model parameters, expectation maximization algorithm, protein secondary structure prediction, bioinformatics
Conference Title : ICSRD 2020 : International Conference on Scientific Research and Development
Conference Location : Chicago, United States
Conference Dates : December 12-13, 2020