

Analysis and Comparison of the Kernel Accuracy for Saccharomyces Genus Protein Sequence Classification

Hyo Jung Chun, Dong Woo Hong, Taeseon Yoon

Hankuk Academy of Foreign Studies, Yongin, Gyeonggi, South Korea
 hjchun96@gmail.com, hdongwoo1479@gmail.com, tsyoon@hafs.hs.kr

Abstract—Even as the modern protein sequencing technology is uncovering great amount of new protein, such data are futile without knowledge of the functions the protein sequences encode. To overcome the limitation of experimental analysis, utilizing Support Vector Machine and kernel methods for functional prediction of unannotated protein has become a promising topic of research in the field of computational biology, inducing many researchers to develop kernels with improved accuracy and efficiency. In this paper, we assigned the Gaussian, Polynomial, and Normal kernels to each Sensus Stricto, Sensus Lato, and Petite-Negative groups of the Saccharomyces fungus species, for which the kernel showed the greatest accuracy in protein sequence classification. From the result we discovered the sequential shapes of the proteins and detected similarities of the structural linearity among the proteins species belonging in the same group. The resulting data allow us to provide an important categorization of kernels that will predict protein function with the greatest accuracy depending on the group of the Saccharomyces the protein belongs for future researches using sequential analysis for prediction.

Keywords—Support Vector Machine, Kernel methods, Protein sequence, Function prediction



Hyo Jung Chun was born in Seongnam, South Korea in 1996.6.04. This author has previously studied abroad in Boulder, Colorado and is currently studying at Hankuk Academy of Foreign Studies in Yongin, Gyeonggi. This author has participated in another conference related to computational science and wishes to study artificial intelligence, robotics, and business in the future.



Dongwoo Hong was born in Busan, South Korea, in 1996.11.14. He is currently studying at Hankuk Academy of Foreign Studies in Yongin, Gyeonggi, South Korea. He wants to major molecular biology and biotechnology in the future.



Taeseon Yoon was born in Seoul, Korea, in 1972. He was Ph.D. Candidate degree in Computer education from the Korea University, Seoul, Korea, in 2003. From 1998 to 2003, he was with EJB analyst and SCJP. From 2003 to 2004, he joined the Department of Computer Education, University of Korea, as a Lecturer and Ansan University, as an Adjunct professor. Since December 2004, he has been with the Hankuk Academy of Foreign Studies in Yongin, Gyeonggi, where he is a Computer Science and Statistics Teacher. He was the recipient of the Best Teacher Award of the Science Conference, Gyeonggi-do, Korea, 2013.