

Analysis of HSV-1 and HSV-2 that Cause Herpes Simplex with Apriori Algorithm, Decision Tree, and Support Vector Machine

Juhyun Lee *, Yeojin Jung *, Seungho Shin*, and Taeseon Yoon**

*International course, HAFS (Hankuk Academy of Foreign Studies), Republic of Korea

** Faculty in HAFS, Republic of Korea

juhyun62015@gmail.com, alwaysujj@gmail.com, leoshin99@naver.com, tsyoon@hafs.hs.kr

Abstract— The Herpes Simplex is a common fluid-intermittent disease that is common among not only adults but also children. Several symptoms include blisters and sores forming on infection sites that are mainly located in oral and genital body parts. The two main types of viral infection, commonly referred to as HSV-1 and HSV- 2, are both part of the same Herpes Simplex Virus group (HSV), but they exhibit different characteristics in the main target of infection and the location of the infection sites. Normal cases of HSV-1 cause oral lesions and has less dangers of cross-generation infection when the female is pregnant, whereas HSV-2 cause genital lesions and usually is co-infected with human immunodeficiency virus type 1(HIV-1). Currently, no specific permanent vaccination is available for this disease and since it is chronic and can be transmitted as long as symptoms are present, the development of vaccination is significant at this stage. In order to provide further possibilities for vaccination development, we analyzed the genomes of HSV-1 and HSV-2 by using Apriori algorithm, Decision tree algorithm, and Support Vector Machine. Thus, we extracted results that showed the similarity and differences of the two genomes, mainly focusing on the prominent distribution of amino acids in each of the 3 categories we have chosen – dividing the genomes up in to 13, 17, and 19 windows. Through this, we have showed further potential in developing effective vaccination more specific to each type of Herpes Simplex.

Keyword— Apriori algorithm, Decision Tree, Herpes Simplex, HSV-1(herpes simplex virus type 1), HSV-2(herpes simplex virus type 2), SVM(Support Vector Machine)



Juhyun Lee was born in Republic of Korea in 1999. She is currently attending HAFS. She is passionate in learning chemistry, biology, and pharmacology. This is her first trial to write a paper. Although she faced some difficulties when writing the paper, she didn't give up and kept working hard. She wants to develop a new medicine in the future, especially a medicine that can cure cancer.



Yeojin Jung is currently a junior at HAFS. She is interested in a variety of fields including bioinformatics, neuroscience, marketing, and economics, and hopes to merge these fields together in to a new area of study in the future. Through her unique perspective in viewing things around her and original thoughts in linking ideas of different fields, she is currently trying hard to accumulate expertise in a variety of fields as to accomplish her goal.



Seungho Shin was born in Seoul, Korea, in 1972. Since 2015, he has been studying in the International Department, Hankuk Academy of Foreign Studies. He is especially interested in the field of bioinformatics, and he wishes to study artificial intelligence and robotics in the future.



Taeseon Yoon was born in Seoul, Korea, in 1972. He received his Ph.D. 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, where he was a computer science and statistics teacher. He was the recipient of the Best Teacher Award of the Science Conference, Gyeonggi-do, Korea, 2013.