

Design of Educational Big Data Application using Spark

Min-Seon Lee*, Dong-Ryeol Shin

Department of Information and Communication Engineering, Sungkyunkwan University

2066, Seobu-ro, Jangan-gu, Suwon-si, Gyeonggi-do, Korea

mslee01@skku.edu, drshin@skku.edu

Abstract—Educational institutions utilize various academic systems. Systems related to various tasks such as academic administration, general administration, and subsidiary institution administration are used. These systems provide information with complex and diverse attributes. Students and faculty want to get the information quickly. Also, systems are important for providing services for students and faculty to use their academic information.

The academic information has large datasets, among which lecture information can be used to find students' learning patterns. So we use an FP-Growth algorithm that compresses data of frequent items into a frequent pattern tree and then divides the compressed data into a set of related condition data for one frequent item and mining it separately does.

Keyword— Spark, MLlib, Recommendation Algorithm, Big Data



Min Seon Lee received the B.E degree in the Department of Computer Engineering from Sungkyunkwan University, Korea, in 2008. Currently, he is working for M.E. in Department of Computer Science and Engineering at Sungkyunkwan University, Korea. His research interests include Machine learning and Big data Platform.



Dong Ryeol Shin received the B.S., M.S., and Ph.D. degrees in Electrical Engineering from the Sungkyunkwan University in 1980, the Korea Advanced Institute of Science and Technology (KAIST) in 1982, and the Georgia Institute of Technology in 1992, respectively. During 1992-1994, he had worked for Samsung Data Systems, Korea, where he was involved in the research of Intelligent Transportation Systems. Since 1994, he has been with the Department of Computer Science and Engineering at Sungkyunkwan University where he is currently a Professor in Network Research Group. His current research interests lie in the areas of mobile network, ubiquitous computing, cloud computing, and bioinformatics. And he is actively involved in the security of vehicular area networks, and the implementation and analysis of big data platform, applicable to 3D image processing of robotic arms.