

강의개요

Dimensionality Reduction

바이오 데이터는 고차원의 데이터로 구성되어 있는 경우가 많다. 일례로 유전자 발현 데이터를 생각하면, 일반적으로 유전자가 변수(feature)가 되기에 인간의 경우 20,000 차원의 이상으로 구성된 데이터라 할 수 있으며 이러한 고차원의 데이터를 이용하여 분석을 진행하는 것은 계산학적으로도 간단한 일이 아니다.

본 강의에서는 이러한 고차원 데이터를 분석할 때 활용할 수 있는 dimensionality reduction (차원 축소) 개념과 대표적인 방법들, 활용 방안에 대해 설명한다. 가장 대표적인 방법인 PCA (principal component analysis)를 비롯하여 ICA (independent component analysis), NMF (non-negative matrix factorization) 등 여러 방법론에 대해 간단히 소개한다. 또한 deep neural networks을 이용하여 dimensionality reduction 목적으로 사용할 수 있는 autoencoder의 개념과 활용에 대해서도 설명한다. 각 방법들에 대한 기본 개념 소개와 함께 실제 데이터를 이용한 활용 예시를 통해 이러한 방법들이 바이오 데이터 분석에 어떻게 활용될 수 있는지 보인다.

본 강의에서는 수학적인 부분을 아예 제외할 수는 없으나, 각 방법의 특성을 비롯한 주요 개념 이해를 목적으로 설명한다. 또한 이론 위주의 강의이나 R이나 python을 이용하여 해당 방법들을 어떻게 활용할 수 있는지에 대해서도 간단히 소개한다. 직접적인 R이나 python 실습이 진행되지는 않지만, 강의 중 소개되는 코드 이해를 위해서는 R과 python 기초 문법을 사전에 알고 있으면 도움이 될 수 있다.

*참고강의교재:

강의 자료

*교육생준비물:

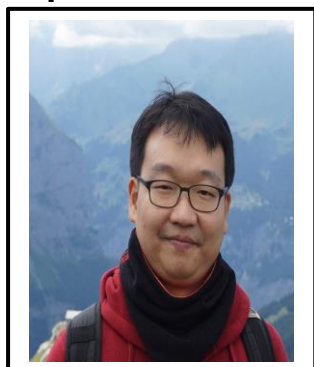
강의 수강을 위한 개인 컴퓨터

* 강의 난이도: 초급

* 강의: 이제근 교수 (송실대학교 의생명시스템학부)

Curriculum Vitae

Speaker Name: Je-Keun Rhee, Ph.D.



► Personal Info

Name Je-Keun Rhee
Title Assistant Professor
Affiliation Soongsil University

► Contact Information

369 Sang-doro, Sangdo-dong, Dongjak-gu, Seoul
Email jkrhee@ssu.ac.kr
Phone Number 02-828-7038

Research interest : Cancer Genomics, Machine Learning

Educational Experience

2004 B.S. in Life Science, Korea University, Korea
2004 B.S. in Computer Science & Engineering, Korea University, Korea (Double Major)
2014 Ph.D. in Bioinformatics, Seoul National University, Korea

Professional Experience

2011 Visiting Scholar, School of Informatics and Computing (SolC), Indiana University, USA
2014-2018 Research Professor, Cancer Research Institute / Department of Medical Informatics, The Catholic University of Korea, Korea
2018-2019 Assistant Professor, School of Dentistry, Pusan National University, Korea
2019- Assistant Professor, School of Systems Biomedical Sciences, Soongsil University, Korea

Selected Publications (5 maximum)

1. Bonil Koo, Je-Keun Rhee, Prediction of tumor purity from gene expression data using machine learning, Briefings in Bioinformatics, 22(6):bbab163, 2021. (Corresponding author)
2. Yeongjoo Kim, Ji Wan Kang, Junho Kang, Eun Jung Kwon, Mihyang Ha, Yoon Kyoung Kim, Hansong Lee, Je-Keun Rhee, Yun Hak Kim, Novel deep learning-based survival prediction for oral cancer by analyzing tumor-infiltrating lymphocyte profiles through CIBERSORT, OncoImmunology, 10(1):e1904573, 2021 (Co-corresponding author)
3. Joong Chae Na, Inbok Lee, Je-Keun Rhee, and Soo-Yong Shin, Fast single individual haplotyping method using GPGPU, Computers in Biology and Medicine, 113:103421, 2019. (Co-corresponding author)

4. Je-Keun Rhee, Soo-Jin Kim, Byoung-Tak Zhang, Identifying DNA Methylation Modules Associated with a Cancer by Probabilistic Evolutionary Learning , IEEE Computational Intelligence Magazine, 13(3): 12-19, 2018. (First author)
5. Joong Chae Na, Jong-Chan Lee, Je-Keun Rhee, and Soo-Yong Shin, PEATH: Single Individual Haplotyping by Probabilistic Evolutionary Algorithm with Toggling, Bioinformatics, 34(11): 1801-1807, 2018. (Co-corresponding author)