

강의개요

SVM (Support Vector Machine)의 개념과 활용

SVM은 분류 문제에 해결함에 있어 딥신경망, 그래디언트 부스팅과 함께 가장 좋은 성능을 보이는 감독 학습 기법으로, 기계학습의 아버지라 불리는 Vladimir Vapnik 교수님의 약 30년간의 기계학습 연구 끝에 탄생한 역작이다. SVM은 기본적으로 큰 마진 (large margin)을 갖는 선형 분류기를 학습하며, 커널 (kernel)이라는 매커니즘을 이용하여 원래의 데이터 공간 대신 소위 피쳐 공간 (feature space)에서 예측에 효과적인 선형 분류기를 학습할 수 있다. 이와 같이, 커널을 이용하는 SVM은 딥신경망과 마찬가지로 데이터의 새로운 표현형을 학습할 수 있기 때문에 비선형 결정 경계를 갖는 분류 문제 해결에 효과적이다.

SVM은 벡터 형 데이터 뿐만 아니라 시퀀스, 그래프 등의 구조를 갖는 데이터에 대해서도 커널을 정의함으로써 적용 가능하며, 다른 기계학습 기법에 비해 과적합 우려가 적으며, 특히 딥신경망 학습과 비교해서는 학습 문제가 볼록 최적화로 귀결되기 때문에 최적의 학습 파라미터를 찾기가 훨씬 용이하다는 장점을 지닌다. 특히, 후자의 사실은 분류기를 구성하는 학습 파라미터의 기여도를 통해 중요 인자를 추출함에 있어 큰 장점으로 작용할 수 있다.

본 강의에서는 SVM의 개념과 학습 최적화 문제, 듀얼 최적화 문제, 커널 SVM에 대해 이해하고, 이를 바탕으로 바이오 데이터를 활용한 문제 해결에 적용할 수 있는 역량을 함양한다.

*교육생준비물:

노트북 (메모리 8GB 이상, 디스크 여유공간 30GB 이상)

구글 계정 생성

Anaconda 다운로드 및 설치 (<https://www.anaconda.com/>)

* 강의 난이도: 초급

* 강의: 이상근 교수 (고려대학교 정보보호학과)

Curriculum Vitae

Speaker Name: Sangkyun Lee, Ph.D.



► Personal Info

Name Sangkyun Lee
Title Assistant professor
Affiliation Korea University

► Contact Information

Address 45, Anam-ro, Seongbuk-gu, Seoul, Republic of Korea
Email sangkyun@korea.ac.kr
Phone Number 02-3290-4890

Research interest : Machine Learning, Secure AI, Collaborative Learning

Educational Experience

2003 B.S., Seoul National University
2005 M.S., Seoul National University
2011 Ph.D., University of Wisconsin-Madison, USA

Professional Experience

2011-2014 Post-doc researcher, Collaborative Research Center (SFB876), TU Dortmund University, Germany
2015-2017 Principal Investigator, Collaborative Research Center (SFB876), TU Dortmund University, Germany
2017-2019 Assistant Professor, Department of Computer Science, Hanyang University ERICA
2020-Current Assistant Professor, Department of Cybersecurity, Korea University

Selected Publications (5 maximum)

1. Data Quality Measures and Efficient Evaluation Algorithms for Large-Scale High-Dimensional Data, Hyeongmin Cho and Sangkyun Lee, Applied Sciences, 2021
2. Sparse Portfolio Selection via the sorted ℓ_1 - Norm, Philipp J. Kremer*, Sangkyun Lee*, Małgorzata Bogdan, and Sandra Paterlini, Journal of Banking & Finance, 2020
3. Structure Learning of Gaussian Markov Random Fields with False Discovery Rate Control, Sangkyun Lee, Piotr Sobczyk and Małgorzata Bogdan, Symmetry, 2019
4. Compressed Learning of Deep Neural Networks for OpenCL-Capable Embedded Systems, Sangkyun Lee and Jeonghyun Lee, Applied Sciences, 2019

5. The mutational landscape of MYCN, Lin28b and ALKF1174L driven murine neuroblastoma mimics human disease, Bram De Wilde, Anneleen Beckers, Sven Lindner, Althoff Kristina, Katleen De Preter, Pauline Depuydt, Pieter Mestdag, Tom Sante, Steve Lefever, Falk Hertwig, Zhiyu Peng, Le-ming Shi, Sangkyun Lee, Elie Vandermarliere, Lennart Martens, Björn Menten, Alexander Schramm, Matthias Fischer, Johannes Schulte, Jo Vandesompele and Frank Speleman, Oncotarget, 2017