



2018년도 제1회 통계적 방법론 공개 강좌

과목 소개

Decision Trees and Random Forests

Decision trees는 재귀적으로 데이터를 분할하여 만든 계층적, 축차적 나무구조 모형으로 데이터의 내재된 정보를 의미 있게 시각적으로 표현하는 방법이다. 예측 변수의 확률 공간을 분할하여 반응 값을 예측하고 복잡한 데이터 구조를 쉽게 해석하기 위해 주로 이용된다.

Random forests는 수많은 decision trees로 구성되어 각 모형의 결과를 종합하는 분류 방법이다. 나무구조 모형의 해석 용이성을 희생함에도 분할 변수를 랜덤하게 선택하는 독특한 방식으로 높은 예측력을 가진다. 본 강좌에서는 기본적인 이론과 주요 알고리즘, 그리고 실제 데이터 분석에 활용 사례를 소개한다.

고 려 대 학 교 통 계 연 구 소

<http://stat.korea.ac.kr/>