**연세-네이버클라우드 데이터사이언스 교육과정**

과목명: 의사결정나무

데이터: Deafult.csv (ISLR)

과제 설명:

R 프로그램의 ISLR 패키지에 포함되어 있는 Default 데이터는 세 개의 설명변수(학생 여부, 부채, 소득)에 따른 사람들의 채무 불이행 여부에 대한 데이터이다.

1. Default 데이터를 Train 데이터와 Test 데이터로 분할 하시오.
2. Train 데이터를 사용하여 옵션을 사용하지 않은 Default Tree를 적합 하시오.
3. Hyper-Parameter Tuning을 통해 Best Tree를 적합 하시오.
4. Default Tree와 Best Tree의 성능을 ROC 곡선을 이용하여 비교 하시오.
5. (model).feature\_importances\_ 명령어를 사용하여 두 의사결정나무의 변수 중요도를 비교 하시오.