**R을 이용한 통계분석 – Scorecard**

**실습 과제 Description**

* 목표
  + German credit data (R 오브젝트명 : german.RData)를 사용하여 scorecard 모형을 개발함
  + Scorecard를 엑셀 형태로 정리함
  + Scorecard 도출 후, validation 데이터를 활용한 K-S 통계량을 산출함
* 분석 과정 및 분석요청 사항
  + Sampling
    - Training (learning) sample을 전체 데이터 중 70%, 나머지는 validation data로 사용함
  + Data 시각화에 기반한 자료탐색
    - R 패키지 ggplot2를 활용한 주요 변수 도출
  + Classing 작업
    - 중요변수에 대한 classing 작업 수행
    - 작업 후, 각 attribute에 대한 구성비 및 WOE 산출/패턴분석
  + Logistic regression의 적용
  + Scaling 작업 적용 및 scorecard 도출
  + Scorecard에 대한 assessment
    - Validation 데이터로 K-S(Kolmogorov-Smirnov) 통계량 산출
* 과제 평가 기준 (100%)
  + Scorecard 개발을 위한 과정을 잘 따랐는가? (25%)
  + 옳은 분석방법을 적용하였는가? (20%)
  + 분석결과에 대한 해석이나 결론이 합당한가? (30%)
  + 개발과정을 명확히 표현하였는가? (15%)
  + 최종 scorecard의 성능은 어떠한가? (10%)

이상.