

### III. 팀별 대표 프로젝트 수행 결과

프로젝트 기반 훈련 (핀다팀) 리더 - 김태영

(인터뷰 및 발표 대상 : O/X) 팀원 - 김민환, 신해솔

연번	훈련생명	생년월일	소속(출신학교)	전공	관련경력 또는 자격증명
1	***	*빈칸제출	*빈칸제출	*빈칸제출	*빈칸제출
2	***				
3	***				
4	***				
5	***				

프로젝트 주제 : 한도조회 유저의 이탈 분석

프로젝트 기반 훈련 목표		프로젝트 기반 훈련 내용
핀다의 중요 서비스인 비교대출 서비스에서의 유저 이탈에 영향을 주는 피처를 생성하고, 이를 기반으로 유저 이탈 위험을 예측하는 모형을 구축		<ul style="list-style-type: none"> <li>- SPARK SQL 기반 ETL 작업 능력 숙련을 통해 기업 내부 데이터 마트를 활용하여 생존 분석용 데이터 처리</li> <li>- 직접 구축한 테이블의 전처리 및 이상치 제거를 통해 모형 학습의 완성도 향상</li> <li>- 기존 생존 분석 모형을 변형 및 응용하여 기업의 실정에 적합한 모형 구축</li> </ul>
프로젝트 단계	단계별 세부 내용	
1. 생존분석 데이터 적재	<ul style="list-style-type: none"> <li>- 기업의 내부 데이터 마트를 활용하여 변수 탐색 및 테이블 결합을 통한 새로운 변수 생성</li> <li>- 생존분석 실행이 가능하도록 변수 처리 및 테이블 생성</li> </ul>	
2. 전처리 및 이상치 제거	<ul style="list-style-type: none"> <li>- 생성된 테이블의 누락값에 대한 분석 및 누락값의 성격을 구분하여 0으로 대체 혹은 마할라노비스 거리법을 이용하여 적절한 결측값 대체 진행</li> <li>- 논리상 부적절하거나 특정 백분위수 기준으로 이상치 제거</li> </ul>	
3. 생존분석 모형 구축 및 학습	<ul style="list-style-type: none"> <li>- 핀다 내 계약을 생존, 그 외 계약을 사망으로 설정하여 생존기간을 예측 (모형) cox propotional hazard model, RandomSurvivalForest model</li> </ul>	
4. 이탈 분석 영향 feature 및 이탈 Alert 모형 구축	<ul style="list-style-type: none"> <li>- 생존분석 결과 생존기간을 짧게하는(즉, 타사로 이탈하여 계약하는) feature의 중요도 및 부호(양/음) 분석</li> <li>- Shapley Value를 도입하여 유저가 가진 feature값이 생존기간에 어떤 영향을 주는지 밝히는 모형을 도입하여 유저의 특정 feature 값이 변할 경우 이탈의 위험이 이탈까지 남은 기간으로 도출되도록 하여 이탈 유저에 대한 새로운 관리 가능성을 살펴봄</li> </ul>	
융합여부 및 내용		<ul style="list-style-type: none"> <li>◦ 융합여부 : -</li> <li>◦ 융합내용 : -</li> </ul>