**문제해결을 위한 빅데이터 활용\_프로젝트 기획안**

기획안 작성일자 : 2021-08-12

| 팀 명 | 2조 : 카드긁조 |
| --- | --- |
| 프로젝트  주제 및 개요 | 가계 대출 연체자 예측 모델 개발 |
| 프로젝트  수행 방향 | 1. EDA  - 대출 신청자료/신용평가사 data/ 이전 대출 기록 등의 feature 탐색  - 데이터 전처리  - data간 통합 처리  2. Model 학습/예측/평가  - 차원 축소, 하이퍼 파라미터 튜닝  - 예측 모델 학습  3. 최종 모델 도출  - 최적 모델 세팅  4. 인사이트 도출  - 신용대출 서류 스크리닝  - 고객 신용정보에 따른 고객 segmentation  - segmentation에 따른 최적화된 금융 상품(차 보험, 차/집 담보 대출) 제안  - 연령대별,기혼/미혼,자녀둔가족 특성 분석  ● 데이터소개 및 기타  - data : Kaggle Home Credit Default Risk  (출처: <https://www.kaggle.com/c/home-credit-default-risk>)  - 알고리즘: XGBoost, LigthGBM, Voting, Averaging, Stacking, Catboost, SVM 등  - 환경: Python, AWS |
| 프로젝트 조직  (구성원 및 역할) | ● 역할분담   | 팀장 | : 최두영 | | --- | --- | | 팀원1 | : 김장호 | | 팀원2 | : 복혜빈 | | 팀원3 | : 이성준 | |
| 프로젝트 추진 일정 | -● 일정  8/05 ~ 8/09: 주제 선정 및 일정 수립  8/09 ~ 8/13: EDA  8/13 ~ 8/17: Model 학습/예측/평가 및 최종 모델 도출  8/18 ~ 8/19: 인사이트 도출 및 발표 준비  8/20 : 발표 |

참고차 아래 넣어둘게요 내용은 합의 필요! (해당 문구 삭제 예정)

| **1** | **프로젝트 구상 및 착수** |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **1.1** | **주제선정** | **사람 이름** | **21년 8월 5일** | **21년 8월 6일** | **1** |
| 1.1.1 | 기획안작성 |  | 21년 8월 10일 | 21년 8월 12일 | 2 |
| **2** | **데이터 수집** |  |  |  |  |
|  |  | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 | 4 |
|  |  |  | 21년 8월 15일 | 21년 8월 15일 |  |
|  |  | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 | 3 |
|  |  |  | 21년 8월 15일 | 21년 8월 15일 | 0 |
| **3** | **데이터 전처리** |  |  |  |  |
| 3.1 |  | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 | 0 |
| 3.2 |  | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 | 0 |
| 3.2.1 |  | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 | 0 |
| 3.2.2 |  | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 | 0 |
| 3.3 |  |  | 21년 8월 15일 | 21년 8월 15일 | 0 |
| 3.3.1 |  |  | 21년 8월 15일 | 21년 8월 15일 | 0 |
| **4** | **데이터 분석 및 시각화** |  | | 2.1 |  | | --- | --- | |  |  | | 2.2 | 주제 관련 인사이트 도출 | | 2.3 |  | |  |  |
| 4.1 |  | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 | 0 |
| 4.2 |  | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 | 0 |
| 4.3 |  | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 | 0 |
| 4.4 |  | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 | 0 |
| **4** | **모델링 (모형 구현)** |  |  |  |  |
| **4.1** | **모델설정** | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 | 0 |
| 4.1.1 | Test 모델 구축 | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 | 0 |
| 4.1.2 | 모델 학습 | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 | 0 |
| 4.1.3 | 모델 비교 및 선정 | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 | 0 |
| **4.2** | **모델 학습 및 구현** |  | 21년 8월 15일 | 21년 8월 15일 |  |
| 4.2.1 | 입력 데이터 테스트 | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 |  |
| 4.2.2 | Hyper parameter 테스트 | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 |  |
| **4.3** | **모델 적용 및 평가** |  | 21년 8월 15일 | 21년 8월 15일 |  |
| 4.3.1 | proto type 모델 학습 | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 | 0 |
| 4.3.2 | 모델 적용 | 사람 이름 | 21년 8월 15일 | 21년 8월 15일 | 0 |