**IM Digital Academy 3기 최종프로젝트**

|  |  |
| --- | --- |
| 팀명 | 알파코 채무불이행 감사팀 |
| 주제명 | 채무불이행 예측 모델 구축 및 리스크 관리 전략 제안 |

1. **역할분담**

|  |  |
| --- | --- |
| 이름 | 역할 |
| 구경서 | 프로젝트 관리, 연구배경 및 도메인 조사 |
| 김자연 | EDA, 모델링, 최종 발표 |
| 박훈석 | 기획안 작성, EDA, 모델링 |
| 여혁수 | EDA, 모델링, 결과 분석 |
| 현지훈 | EDA, 모델링, 발표자료 기획 |

**2. 주제 선정 배경**

최근 지속적인 경기 불안과 금리 인상으로 인해 가계부채 리스크가 금융기관의 큰 이슈로 떠오르고 있습니다. 특히 채무불이행은 금융기관의 재무 건전성을 위협할 뿐 아니라 사회경제적 문제로 확대될 수 있습니다. 개인의 신용도 하락과 경제적 어려움을 가중시킬 수 있어 이를 사전에 예측하고 방지하는 기술의 필요성 또한 부각되고 있습니다. 따라서 인위적으로 생성된 금융 데이터를 바탕으로 머신러닝을 활용한 채무불이행 예측 모델을 구축하여, 금융기관이 선제적으로 리스크를 관리할 수 있는 전략적 도구를 마련하는 것이 중요하다고 판단하였습니다.

효과적인 채무 불이행 여부 예측 모델은 금융 안정성을 높이는 데 기여할 뿐만 아니라, 금융 서비스 제공자와 고객 간의 신뢰를 강화하는 데 중요한 역할을 합니다. 이러한 기술은 금융 기관이 보다 신중하고 효율적인 의사결정을 내릴 수 있도록 돕고, 고객들에게는 더 나은 금융 서비스를 제공할 수 있는 기반을 마련해줍니다.

**3. 프로젝트 목표**

이번 프로젝트는 다양한 금융 데이터를 활용하여 채무 불이행 가능성을 예측하는 AI 알고리즘을 개발하는 것을 목표로 합니다. 데이터 분석과 인공지능 기술을 활용하여 금융 생태계를 더욱 안전하고 신뢰할 수 있는 방향으로 이끌고, 이를 통해 지속 가능한 금융 환경 조성에 기여하고자 합니다.

* 머신러닝 기반 채무불이행 예측 모델 구축 및 성능 최적화

다양한 금융 데이터를 활용하여 채무 불이행 가능성을 예측하는 AI 알고리즘을 개발합니다. 이를 통해 금융 기관이 채무 불이행 리스크를 사전에 식별하고 적절한 대응 전략을 수립할 수 있도록 지원합니다.

* 채무불이행과 관련된 주요 특징(feature) 분석 및 발굴

데이터 전처리, 피쳐 엔지니어링, 차원 축소 등의 기술을 적용합니다. 또한 데이터의 불균형 문제를 해결하고, 예측 정확도와 정밀도를 향상시킵니다.

* 채무불이행 가능성이 높은 고객을 미리 식별하여 금융기관의 리스크 관리 전략을 제시

**4. 활용 데이터**

* DACON 채무불이행 예측 대회 데이터셋
  + 주요 컬럼: UID, 주거형태, 연간소득, 현재직장근속연수, 체납세금압류횟수, 개설된신용계좌수, 신용거래연수, 최대신용한도, 신용문제발생횟수, 마지막연체이후경과일수, 개인파산횟수, 대출목적, 대출상환기간, 현재대출잔액, 현재미상환신용액, 월상환부채액, 신용점수, 채무불이행여부(타겟 변수) 등

### **5. 분석 방안**

* 탐색적 데이터 분석(EDA) 수행 및 데이터 특징 파악
* 결측치, 이상치 처리 및 변수 선택과 Feature Engineering 수행
* XGBoost, LightGBM, Random Forest 등 머신러닝 모델 구축 및 최적화
* 교차검증(Cross-validation) 및 하이퍼파라미터 튜닝을 통해 모델 성능 극대화
* 주요 변수 중요도 분석을 통해 채무불이행에 가장 큰 영향을 주는 요인 도출

**6. 결론**

이번 프로젝트는 금융 데이터를 활용하여 채무 불이행 여부를 예측하는 AI 알고리즘을 개발하는 데 초점을 맞추었습니다. 이를 통해 금융기관이 채무 불이행 리스크를 사전에 식별하고, 보다 효율적인 대출 관리 및 의사결정을 지원할 수 있는 기반을 다져볼까 합니다.

데이터 분석을 통해 주요 변수(신용 점수, 월 상환 부채액, 체납 세금 압류 횟수 등)를 식별하고, 머신러닝 모델을 활용하여 예측 성능을 평가합니다. 또한, 결과를 분석하여 채무 불이행 위험이 높은 집단의 특성을 도출하며 실질적인 인사이트를 내보려 합니다.

본 프로젝트는 팀의 창의적 접근과 데이터 분석 능력을 통해 금융 생태계의 안정성을 높이는 데 기여할 가능성을 보여주고자하며, 이를 통해 더 안전하고 신뢰할 수 있는 금융 환경 조성에 한 걸음 더 다가갈 수 있을 것입니다.

### **기대 효과**

* 머신러닝 기반의 정확한 채무불이행 예측으로 금융기관의 부실 위험 사전 대응 가능
* 금융 데이터 분석 역량 증대 및 금융권 취업을 위한 데이터 분석 역량 입증
* 실무적 관점에서 금융기관이 선호할 수 있는 현실적인 데이터 분석 보고서 제공

### **미래 과제 및 방향**

* 실제 금융기관의 실제 데이터와 비교 분석하여 모델의 현실성 향상
* 설명 가능한 AI(Explainable AI)를 적용하여 신용평가 시스템의 신뢰성 제고