

주요 분석 내용

이 문서는 다음과 같은 내용을 포함하고 있습니다:

1 시스템 개요

- 저축은행 신용평가를 위한 AI 기반 예측 시스템
- XGBoost, DNN, Random Forest, TabNet 등 다양한 모델 지원
- 앙상블 기법을 통한 정확도 향상

2 핵심 컴포넌트

- **launcher.py**: 배치 자동 실행 시스템
- **Data_preprocessing**: 데이터 전처리 파이프라인
- **Learning**: 분산 학습 시스템
- **Ensemble**: 대규모 앙상블 처리 (100만회 × 30개 모델)
- **RESTful API**: Django 기반 API 서버

3 시스템 아키텍처

- Dask Distributed 기반 분산 컴퓨팅
- GPU 4개 활용 (NVIDIA TITAN)
- MariaDB 기반 배치 관리
- SFTP를 통한 모델 저장/로드

4 주요 특징

- 자동화된 배치 처리 ($I \rightarrow U \rightarrow C/E$ 상태 관리)
- 374개 베이스 모델을 활용한 앙상블
- 데이터 불균형 처리 (SMOTE, Undersampling)
- 베이지안 최적화를 통한 하이퍼파라미터 튜닝

5 실용 정보

- 설치 방법 및 실행 가이드
- 데이터베이스 스키마 설명
- 사용 예시 및 운영 정보

이 시스템은 신용평가 분야에서 머신러닝/딥러닝을 활용한 종합적인 예측 플랫폼으로, 대규모 병렬 처리와 앙상블 기법을 통해 높은 정확도를 달성하도록 설계되어 있습니다.