K-SW 스퀘어 프로젝트 실적 및 계획

◈ 프로젝트명 : 빅데이터 ‘

◈ 팀명 : XLargeKSQ

◈ 참가자 : 강준후, 김나원, 배지훈, 백승근, 장유진

|  |  |
| --- | --- |
| 9주차  실적  (11.06~11.11) | O 주간 실적 작성   1. LightGBM Parameter Tuning  * LightGBM 모델의 Parameter를 조정하여 66.9%의 정확도를 기록하였음.  1. Random Forest  * TensorFlow 내의 Random Forest를 사용하는 경우, GPU를 사용함에도 불구하고 4h 30m 가량 소요. * Scikit-learn 내의 Random Forest를 사용하는 경우, 음악을 감상한 메뉴와 관련된 3가지 feature만 사용하였을 때 57.7%의 정확도를 나타냄.  1. Super computer  * 슈퍼컴퓨터 시스템이 Python 2에 최적화되어 있는 관계로, 호환되지 않는 Python 3 라이브러리 및 코드를 직접 수행하기에는 다소 무리가 있음. * TensorFlow-GPU 사용 시, 라이브러리를 개별 컴퓨터에서 빌드하면 더 좋은 성능을 보인다고 조언 받음.   기타. 문화 체험: 농구 경기 관람 |
| 10주차  계획  (11.13~11.18) | O 향후계획 작성   1. LightGBM Data Preprocessing  * 모델 인자 조정만으로는 성능 향상에 한계가 있으므로, 별도의 데이터 전처리를 통해 더 의미있는 데이터를 모델에 제공할 예정 * 예시: 사용자별 재감상 빈도, 많이 재생한 장르 등  1. Random Forest  * 전체 Feature를 넣어 모델의 실제 성능 확인 * 모델 인자 조정 * 1에서 전처리한 데이터를 활용하여 분석 시도  1. Neural Network  * TensorFlow를 활용하여 Neural Network 모델 구현 * Parameter: Network의 깊이, 각 Depth 별 Neuron의 수 등 * 1,2와의 성능 비교 * Categorical Data를 분석할 때 Neural Network가 다른 모델에 비해 일반적으로 정확한 지 여부에 대한 문헌 조사 |

\* 매주 목요일 일과 후 작성 후 [sechs11@iitp.kr](mailto:sechs11@iitp.kr)로 메일 제출