

학부생 연구기회 프로그램 (UROP) 공고

◆ 담당교수 : 전병곤	◆ 연구실명 : 소프트웨어 플랫폼 연구실
◆ UROP 연구 과제명 : 특정 딥러닝 모델에 대한 학습 속도 최적화	
◆ 모집대상 : TensorFlow, PyTorch 등 딥러닝 프레임워크에 익숙한 학부생	
◆ 모집기간 : ~ 2019년 6월 말	

연구 배경

- 딥러닝 모델 학습은 아주 많은 연산량을 요하는 작업이기에 GPU 등의 하드웨어 가속기를 이용하는 것이 보편적이다.
- 새로운 모델, 새로운 가속기를 사용하게 될 때 마다 사용자가 수동으로 학습 속도를 최적화해야 하는 문제가 있다.
- 특정 딥러닝 모델을 선택하여, 이를 직접 최적화하는 연구를 진행한다.
- 이 연구를 바탕으로, 새로운 모델과 하드웨어에 대하여 자동으로 학습 속도를 최적화해 주는 시스템을 만들고자 한다.

연구 내용

- 기존에 존재하는 학습 속도를 자동으로 최적화하는 시스템(Apache TVM, TensorFlow XLA, PyTorch JIT)들이 어떻게 딥러닝 모델 학습을 최적화하는지 알아본다.
- 특정 딥러닝 모델들에 대하여 기존 시스템들이 활용하지 못하고 있는 최적화가 있는지 살펴보고, 해당 최적화를 직접 구현해 성능을 향상시킨다.
 - 특정 연산을 위해 최적화된 CUDA kernel 만들기, operation fusion, memory layout optimization 등 다양한 종류의 최적화를 포함함

사전 지식 및 요건

- TensorFlow, PyTorch 등 딥러닝 프레임워크 사용 경험
- [선택] CUDA programming 경험



서울대학교 컴퓨터공학부
Seoul National University
Dept. of Computer Science and Engineering



소프트웨어 플랫폼 연구실
Seoul National University
Software Platform Lab

문의

유경인 (gyeongin@snu.ac.kr)