3월7일~3월13일

3월2주차 졸업논문 주제 선정 및 개념공부

머신러닝과 관련된 edge computing와 Distributed environment 에 대해 공부했습니다.

-edge computing

Edge Computing은 네트워크의 주변부에서 클라이언트 데이터가 가능한 원본 소스에 가깝게 처리되는 분산 정보 기술 (IT) 아키텍처 이다. edge computing으로의 전환은 모바일 컴퓨팅, 컴퓨터 구성 요소의 비용 감소 및 인터넷 사물 (IoT)의 네트워크 장치 수의 증가로 인해 발생한다. Edge Computing은 클라우드 서비스를 대신하여 다운 스트림 데이터 및 IoT 서비스를 대신하여 upstream 데이터에서 네트워크 에지에서 계산을 수행 할 수 있게 하는 기술이다. 구현에 따라, edge computing architecture의 시간에 민감한 데이터는 지능형 장치에 의해 출처에서 처리되거나 클라이언트와 가까운 지리적으로 근접하여 위치하는 중간 서버로 전송될 수 있다. 시간에 덜 민감한 데이터는 과거 분석, 큰 데이터 분석 및 장기 저장을 위해 클라우드로 전송된다. 네트워크를 통해 엄청난 양의 원시 데이터를 전송하면 네트워크 리소스에 엄청난 부하가 발생한다. 경우에 따라 원본 근처에서 데이터를 처리하고 네트워크를 통해 가치가 있는 데이터만 원격 데이터 센터로 보내는 것이 훨씬 더 효율적이다. 또한 edge computing은 지리적으로 분산 된 사용자 기반이 있는 원격 사무실/branch office (ROBO) 환경 및 조직에 도움이 될 수 있다. 이러한 시나리오에서는 중간 규모의 마이크로 데이터 센터 또는 고성능 서버를 원격 위치에 설치하여 클라우드 서비스를 로컬로 복제 할 수 있으므로 성능이 향상되고 장치가 부패하기 쉬운 데이터를 신속하게 처리 할 수 있다. 공급 업체 및 기술 구현에 따라 중개자는 edge gateway, base station, hub, cloudlet 또는 aggregator를 포함한 여러 이름 중 하나로 언급 될 수 있다.

-Distributed environment

분산 환경은 여러 계산 자원이 가상화 기술로 통합된 고성능 컴퓨팅 환경으로 구성되어 있으며, 다수의 사용자가 복잡도가 높은 응용 작업에 대해 지속적인 처리를 요구할 경우, 작업을 효율적으로 분산된 계산 자원들에 해당해야 서비스 처리율을 보장할 수 있다. 사용자가 요구하는 작업의 대부분은 데이터 처리이며, 이것은 데이터의 단순 저장 및 정리 작업부터, 영상 처리, 이미지 처리, 음성 처리 등 정형 및 비정형 데이터의 연산 작업까지 다양하다. 각 작업은 다양한 데이터를 포함하고있으므로 계산 자원에 대한 의존도가 다르다. 데이터의 특성을 고려하지 않고 자원을 할당할 경우 분산 환경에서 작업이 효율적으로 처리될 수 없으며, 작업 처리속도가 느려질 수 있다.