

AI빅데이터 분석 및 스파크 활용 강의계획서

강의	강의명	AI빅데이터 분석 및 스파크 활용
	시간	5시간29분
주 수강대상	SW전공자 및 SW에 관심을 가진 자	
담당교수	성명	조숙경
	소속	인하대학교

강의 기본정보

선수과목 또는 선수학습	해당없음						
강의 성격	본 강의는 AI데이터를 분석하고 스파크를 활용하여 프로그램을 시작할 수 있도록 기본 이해부터 RDD다루기까지 단계적으로 학습한다. 기본적인 기능을 익히고, 한 강좌를 통해 다양한 AI빅데이터 분석 및 스파크를 활용하는 방법을 학습한다.						
강의 목표	<ol style="list-style-type: none"> 1. 빅데이터의 이해한다. 2. 스파크의 개요와 활용방법을 습득한다. 3. 기본 아키텍처, 어플리케이션 등의 기능을 학습한다. 4. 구조적 API 및 기본연산 실행과정을 학습한다. 5. 데이터 베이스 시스템을 이해한다. 6. 스파크 활용방법을 습득한다. 						
강의방법	강의식	토론식	세미나식	실험실습식	인터넷전용	인터넷병행	기타
	✓				✓		

차시별 세부내용

차시	학습주제	학습내용	강의방법	수업방법
1	빅데이터의 이해	HDFS, 맵리듀스, 스파크 탄생 배경	인터넷강의	강의
2	스파크 개요	기능 구성, 하둡 vs. 스파크, RDD 개요, 배치 및 실시간 처리 개요	인터넷강의	강의
3	스파크 설치 및 실행	스파크 설치, 주피터 노트북 설치, 실행 방법	인터넷강의	강의/실습
4	스파크 기본 아키텍처 및 간단 기능 확인	기본 아키텍처, 어플리케이션, 다양한 데이터 프레임, 트랜스포메이션, 액션	인터넷강의	강의/실습
5	구조적 API 및 기본 연산	스키마, 구조적 타입 개요, 구조적 API 실행과정	인터넷강의	강의/실습

6	집계 연산	집계함수, 그룹화, 윈도우 함수, 그룹화	인터넷강의	강의/실습
7	조인	조인 표현식, 조인 타입, 수행 방식	인터넷강의	강의/실습
8	SQL	JDBC/ODBC 연결, SQLite	인터넷강의	강의/실습
9	스파크 SQL	하이브, 스파크 SQL 쿼리 실행 방법	인터넷강의	강의/실습
10	RDD	저수준 API, RDD 다루기	인터넷강의	강의/실습