

교육 과정 소개서.

초격차 패키지 : 한 번에 끝내는 빅데이터처리
with Spark&Hadoop

안내.

해당 교육 과정 소개서는 모든 강의 영상이 촬영하기 전 작성되었습니다.

* 커리큘럼은 촬영 및 편집을 거치며 일부 변경될 수 있으나, 전반적인 강의 내용에는 변동이 없습니다.

아래 각 오픈 일정에 따라 공개됩니다.

- 1차 : 2022년 3월 21일
- 2차 : 2022년 4월 25일
- 3차 : 2022년 5월 23일
- 4차 : 2022년 6월 27일
- 5차 : 2022년 7월 25일

최근 수정일자 2022년 02월 16일



강의정보

강의장	온라인 강의 데스크탑, 노트북, 모바일 등
수강 기간	평생 소장
상세페이지	https://fastcampus.co.kr/data_online_spkhdp
담당	패스트캠퍼스 고객경험혁신팀
강의시간	100시간 예정 (* 사전 판매 중인 강의는 시간이 변경될 수 있습니다.)
문의	고객지원 : 02-501-9396 강의 관련 문의: help.online@fastcampus.co.kr 수료증 및 행정 문의: help@fastcampus.co.kr

강의특징

나만의 속도로	낮이나 새벽이나 내가 원하는 시간대에 나의 스케줄대로 수강
원하는 곳 어디서나	시간을 쪼개 먼 거리를 오가며 오프라인 강의장을 찾을 필요 없이 어디서든 수강
무제한 복습	무엇이든 반복적으로 학습해야 내것이 되기에 이해가 안가는 구간 몇번이고 재생



강의목표

- 빅데이터 분석과 분산 처리의 핵심인 Spark와 Hadoop을 한 번에 마스터하도록 안내해 드립니다.
- 빅데이터분석가나 엔지니어가 실무에서 어떤 일을 하는지 알려드립니다.
- 데이터 분석의 앞 단계인 데이터 파이프라인 구축까지 기술을 확장하도록 안내해 드립니다.

강의요약

- 빅데이터 전문가가 되기 위해 필요한 Spark와 Hadoop 기초 이론부터 확장 라이브러리들까지 빠짐없이 담았습니다. 전매특허 초격차 패키지의 한 번에 끝내는 커리큘럼!
- Spark & Hadoop 설치부터 시작해서 빅데이터 적재, 처리, 분석, 추천 머신러닝까지 실무 빅데이터 처리실습으로 꽉 채운 강의. 총 100시간 강의 중 60시간 실습 파격 구성!
- Python vs. Scala 패캠의 최첨단 추천시스템 알고리즘을 따라 나에게 맞는 프로그래밍 언어를 알아보세요! 물론 둘 다 배우는 것도 가능합니다!

CURRICULUM

01.

꼭 알아야 할 핵심 개념

빅데이터 처리 입문하기
01. 빅데이터 플랫폼의 이해
02. 빅데이터 아키텍처의 이해
03. 데이터 파이프라인 패턴과 오케스트레이션
04. 하둡 클러스터 구축 고려사항
05. 하둡분산파일시스템(HDFS)의 구현원리
06. 하둡맵리듀스(MapReduce)의 이해
Apache Spark 기초
01. Spark 개요(Apache Spark 소개, 특징, 지원언어, Hadoop과 비교, Spark 프로그래밍 모델(RDD), 작동원리, Spark Components 등)
02. Scala와 Python 개요(언어소개, 특징, 변수 선언, 함수 정의 등)
03. What is Spark? (Scala vs. Spark)
04. Why Spark? (Python vs. PySpark)
05. 클러스터 환경 구성(AWS 환경에서 Hadoop 및 Spark 설치)

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

02.

Spark 기초와 핵심

Spark Core
01. Spark RDD 개요(Spark Core Components, Spark Application 배포, RDD Operations 등) RDD 프로그래밍(RDD 생성, RDD Transform, RDD Action 등)
02. Shared Variables(Broadcast, Accumulators)
03. Spark Shell 기반 RDD 기본 프로그래밍
04. Spark Application 빌드 및 배포
Spark SQL
01. Spark SQL 개요(특징, 성능, RDD와 DataFrame/Dataset의 비교, SparkSession) DataFrame/Dataset(생성법, Basic Operations, Queries, Join, 저장, Row Object 처리 등)
02. Catalyst Optimizer 및 Tungsten Project 작동원리
03. Web Notebook 기반 SQL/DataFrame 기본 프로그래밍
04. SparkSQL 기반 항공 데이터 분석
Spark Streaming
01. Spark Streaming(특징 및 기능, DStream의 개념 및 Basic Operations, Stateful Operations, Execution Model, DStream Persistence, Checkpoint, Kafka 연계 등)
02. Sturctured Streaming(DStream의 문제점, Programming Model, Window Operations, Watermarking 등)
03. Twitter 데이터 실시간 수집 및 저장(on Kafka & DW)
Spark ML
01. Spark MLlib(특징, 머신러닝 분류, 데이터 유형, Word2Vec, VectorAssembler, ChiSqSelector 등)
02. MLlib Algorithms(Classification, Regression, Clustering, Collaborative, Filtering, Dimensionality Reduction) ML Pipeline(Estimator, Transformer, Parameter, Pipeline)
03. Model Selection(CrossValidator, TrainValidationSplit) ML Persistence(Status, Predictive Model Markup Language)
04. 한국어 위키 워드 임베딩(Word2Vec)
05. Twitter 랭귀지 분류(KMeans)
06. 손글씨 숫자 분류(MLP)
Graph X

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

03.

Spark&Hadoop Ecosystem

Spark Ecosystem

본 챕터 커리큘럼은 추후 공개 예정입니다.

Hadoop Ecosystem

01. Apache Hive와 Trino

02. 스트림 처리 : Spark Streaming(DStrem), Spark Structured Streaming, Flink

03. 데이터 저장 : HBase, Cassandra, Kudu

04. 워크플로우 관리 : Airflow, Oozie

05. 분산 코디네이션 : Zookeeper

06. 시간 로그 분석 : Elastic Stack(Filebeat, Logstash, ElasticSearch, Kibana)

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

04.

종합 프로젝트

Spark Ecosystem 기반 종합 실습

01. 추천시스템의 개요(추천시스템 아키텍처 및 알고리즘, 배치 추천시스템과 실시간 추천시스템)

02. 음악/아티스트 추천시스템 구현

Hadoop Ecosystem 기반 종합 실습

01. 클릭스트림 분석 시스템 만들기

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

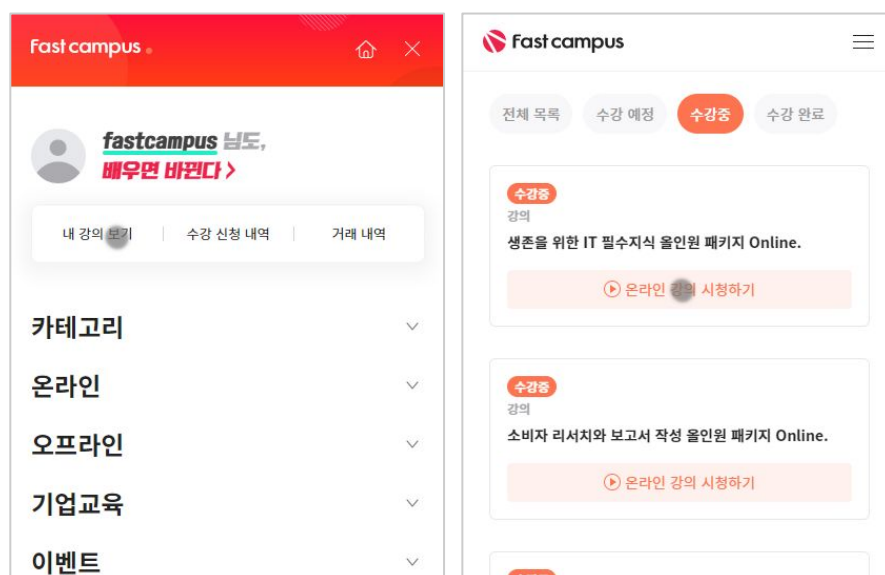


주의 사항

- 상황에 따라 사전 공지 없이 할인이 조기 마감되거나 연장될 수 있습니다.
- 패스트캠퍼스의 모든 온라인 강의는 **아이디 공유를 금지**하고 있으며 1개의 아이디로 여러 명이 수강하실 수 없습니다.
- 별도의 주의사항은 각 강의 상세페이지에서 확인하실 수 있습니다.

수강 방법

- 패스트캠퍼스는 크롬 브라우저에 최적화 되어있습니다.
- 사전 예약 판매 중인 강의의 경우 1차 공개일정에 맞춰 '온라인 강의 시청하기'가 활성화됩니다.



환불 규정

- 온라인 강의는 각 과정 별 '정상 수강기간(유료수강기간)'과 정상 수강기간 이후의 '복습 수강기간(무료수강기간)'으로 구성됩니다.
- 환불금액은 실제 결제금액을 기준으로 계산됩니다.

수강 시작 후 7일 이내	100% 환불 가능 (단, 수강하셨다면 수강 분량만큼 차감)
수강 시작 후 7일 경과	정상(유료) 수강기간 대비 잔여일에 대해 환불규정에 따라 환불 가능

※ 강의별 환불규정이 상이할 수 있으므로 각 강의 상세페이지를 확인해 주세요.