

교육 과정 소개서.

The RED : Hadoop&Spark 아키텍처 딥다이브

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 : 2021년 10월 29일
- 2차 : 2021년 12월 01일

최근 수정일자 2021년 09월 24일



강의정보

강의장	온라인 강의 데스크탑, 노트북, 모바일 등
수강 기간	평생 소장
상세페이지	https://fastcampus.co.kr/data_red_bdl
담당	패스트캠퍼스 고객경험혁신팀
강의시간	5시간 예정 (* 사전 판매 중인 강의는 시간이 변경될 수 있습니다.)
문의	강의 관련 전화 문의: 02-568-9886 수료증 및 행정 문의: 02-501-9396 / help.online@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

- 현존하는 Framework 중 가장 보편적이고 정석으로 사용되는 세 가지 Framework인 Hadoop과 Spark, Kafka 를 학습하여 보다 짧은 시간내에 정확하게 빅 데이터를 처리합니다.

강의요약

- 데이터 엔지니어링 핵심 5가지 Framework 학습
 > 상용되고 있는 Hadoop, Spark, Kafka의 아키텍처는 물론 최근 주목받는 Livy와 Delta Lake까지 모두 학습합니다.
- 이론과 실습을 겸비한 6가지 주제의 Deep Dive
 > 주어진 문제를 해결하기 위해 구체적으로 어떤 툴셋을 사용해야 하는지 Code Walkthrough를 통해 각 Framework 의 feature들을 입증, 정의해봅니다.
- Audi, TOYOTA의 Case Study 까지
 > 글로벌 기업들은 공장 자동화 과정에서 생긴 데이터들을 클라우드 기술과 접목시켜 어떻게 운영하고 있는지 실제 사례를 통해 알아봅니다.



강사

Brandon Lee

약력

- 2020 ~ 현재
Rockwell Automation 실리콘밸리 오피스/ Analytics & ML Architect
- 빅데이터의 분석 및 머신러닝을 할 수 있는 기반 수립
- 자동화 공정 장비 데이터 분석을 통해 향상된 IOT 기반 스마트 제조 모듈 제조에 기여
- 2016 ~ 2020
마이크로소프트 (마운틴뷰) / 수석 데이터 솔루션 아키텍트
- Azure에서 배포, 소비 및 사용을 주도하는 데이터/AI 클라우드 솔루션 설계자
- 고객사의 미션 크리티컬 기업 문제 해결을 위해 전략적 ML 접근 방식을 쉽게 적용하고 제공할 수 있는 아키텍처를 만듦
- 2010 ~ 2016
삼성 SDS (산호세) / 빅데이터 및 클라우드 시니어 매니저
- 삼성전자 빅데이터 및 IoT 클라우드 스택을 최고의 분석 플랫폼으로 전환
- 글로벌 비즈니스 워크로드 운영을 위해 온프레미스 및 클라우드 인프라를 사용한 빅 데이터 및 하이브리드 클라우드 플랫폼 아키텍처 생성
- 2007 ~ 2010
Cimac (산호세) / 소프트웨어 제품 디렉터
- 복잡한 제조 환경을 위한 기계 학습 및 로그 분석 제품 개발
- 프로세스 최적화 및 자동화 솔루션을 구현
- 2006 ~ 2007
Spansion / 선임 R&D 프로그램 매니저
- 장비 가상화 및 프로젝트 엔지니어

CURRICULUM

01.

Hadoop Architecture & File Formats

HDFS: Distributed Data Storage
HDFS Introduction
Hadoop File Formats (Row-based/Column-based)
YARN: Yet Another Resource Manager
Name Node
Data Node
Resource Manager
Node Manager
Shared Edit logs or Journal Node
MapReduce
MapReduce Model
MapReduce Code Example
Zookeeper: Cluster management system
Leader node
Follower node
In-memory Database
Request Processor
Atomic Broadcast

CURRICULUM

02.

Hive LLAP & HBase: Big data NoSQL database

Hive Architecture
Hive Internals
Hive Server
Thrift Client
JDBC Client
ODBC Client
Partitioning & Bucketing
Optimizing Hive Queries
MapReduce concept & code refresher
Hive MapReduce Query Code example
HBase Architecture
HBase Internals
HBase Data Model

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

CURRICULUM

03.

Kafka: Event streaming platform

Event Streaming Platform
App Modernization
Data Modernization
Hybrid/Multi-cloud Streaming
Kafka Internals
Kafka cluster
Partition Replication
Consumer Group
Zookeeper
Confluent Kafka Reference Architecture
Rest Proxy
Schema Registry
KSQL
Kafka Connect
KStreams
Confluent Control Center
Kafka Streams Architecture
Kafka Streams Internals
KSQL Architecture
KSQL Architecture
Reference Architectures
Toyota Lean Manufacturing Across Edge & Azure Cloud
Audi Real Time Data Cluster
Event Streaming Sensor Analytics Pipeline Demo
Event Streaming Sensor Analytics Pipeline Demo

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

CURRICULUM

04.

Spark: Big data processing framework with Streaming batch analytics

Spark Introduction
Spark Introduction
Spark Architecture & Internals
Spark Driver Program
Cluster Manager
Executors
Spark ML Model Development Workflow
Data Exploration
Data Preparation
Model Building & Evaluation
Deployment & Monitoring
History of Spark APIs
Evolving spark APIs to better process data
Spark Session
Spark Session vs. Spark Context
Spark Data Preparation
Data Cleansing & Preparation
Spark ML Pipeline
Spark ML APIs
Spark Cluster Architecture Example
Spark Cluster Architecture Example

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

CURRICULUM

05.

Livy: REST API web service for Spark

Livy introduction
Livy introduction
Livy features
Livy features

CURRICULUM

06.

DeltaLake: ACID compliant and scalable storage to Apache Spark

Deltalake Introduction
Deltalake Introduction
Detalake features
Detalake features

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

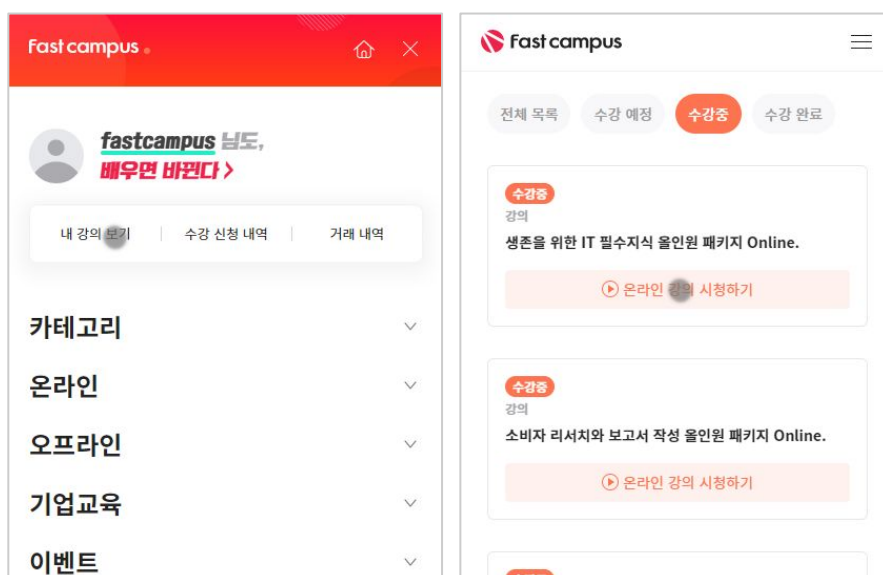


주의 사항

- 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 학원법 환불규정에 따라 환불 가능

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