

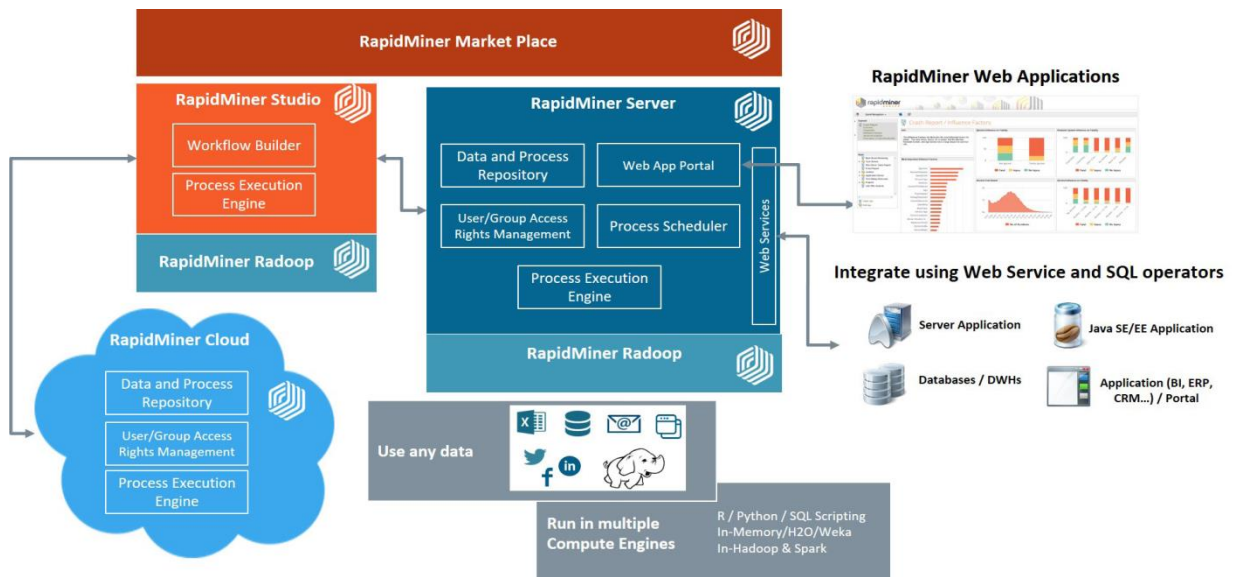
RapidMiner Data Science Platform

A Complete Platform for Predictive Analytics

RapidMiner는 예측 분석을 위한 모델 개발부터 관리까지 완전 GUI 기반으로 작업을 할 수 있는 통합 플랫폼입니다. 다양한 스크립트 언어를 지원 하고 있으며, 분석을 위한 모든 과정을 단일 환경 (Single Suite)에서 관리하기 때문에 업무의 생산성 및 효율성을 증가 시킬 수 있습니다.

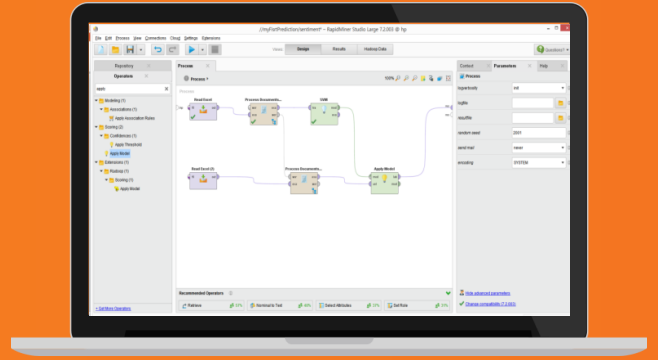
주요 Module

- . RapidMiner Studio – 분석 프로세스를 작업하기 위한 Client
- . RapidMiner Server – 팀 협업 기능, 프로세스 자동 실행 및 스케줄링, 타 시스템과의 연동, web Application 생성
- . RapidMiner Radoop – Big data 환경(Hadoop)에서 쉽고 빠르게 데이터 분석을 할 수 있도록 구성된 모듈



RapidMiner Studio

Data Science Workflow Designer



주요 기능

- Visual Programming 환경
- 강력한 분석 가이드 제공
- 프로세스의 Block화로 재사용 용이
- 1500 이상의 Machine Learning & Data Preparation 함수 제공
- R & Python Scripts와의 연결
- 모든 유형의 데이터에 Access

신속한 데이터 분석

RapidMiner Studio는 정형 데이터 뿐만 아니라 비정형 데이터까지도 신속하게 적재, 결합, 변환 및 분석을 할 수 있는 기능을 제공함으로써 기존 분석 보다 차원 높은 분석을 할 수 있는 환경을 제공합니다.

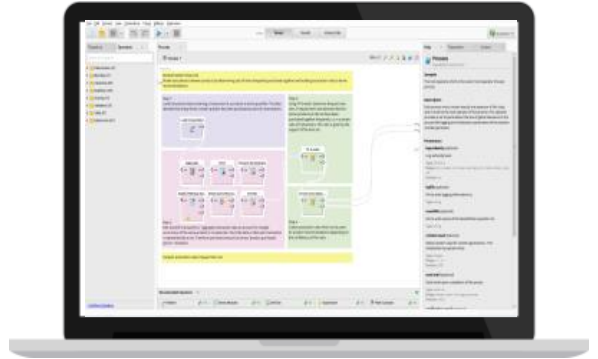
데이터 분석의 생산성 극대화

GUI방식의 RapidMiner Studio는 다양한 분석 가이드 기능들과 기본 제공 템플릿 및 기 정의된 프로세스에 대해서 쉽게 반복 사용할 수 있는 Workflow를 제공합니다.

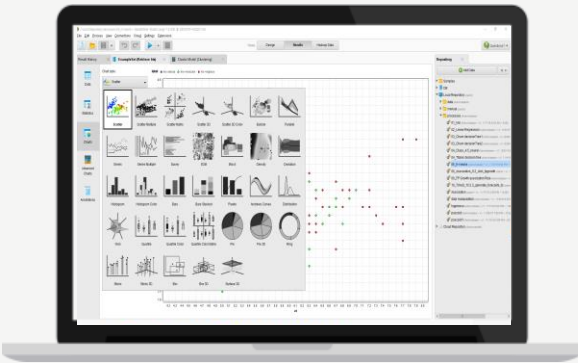
아울러 강력한 알고리즘과 함수로 구성된 1500개 이상의 풍부한 라이브러리와 R, Python 코드와의 쉬운 연계는 모든 use case에 대해 강력한 모델을 구축하는데 도움을 줍니다.

Visual Programming

- 완전 GUI 방식을 통해 여러개의 Operator들을 활용하고 연결하여 모델 적용에 대한 일련의 과정을 코딩 없이 진행
- Process panel을 통한 작업으로 저장 및 재사용이 쉽고, 다른 사람과 분석 과정에 대한 공유가 용이
- 다양한 Application과 연관된 operator 활용을 통해 Advanced 분석이 가능



쉽게 확인 가능한 EDA



- 데이터 탐색을 위한 변수 타입 및 통계량 등을 쉽게 확인할 수 있는 Summary View
- 30가지가 넘는 다양한 기본 차트 및 사용자가 디자인할 수 있는 Advanced 차트
- 데이터의 정보데이터(Meta data)를 기록 및 확인할 수 있는 기능 (Annotation)

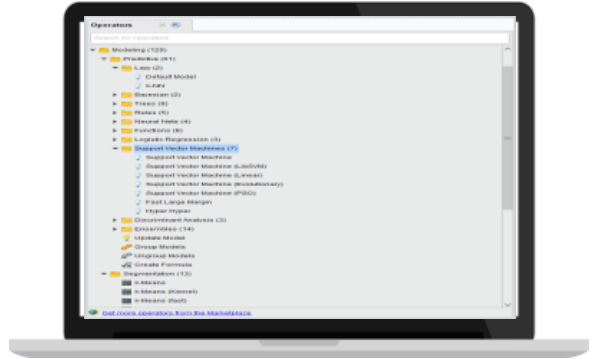
사용자 편의를 위한 Reference 제공

- Rapidminer 홈페이지를 통해 분석하고자 하는 데이터의 가장 적합한 알고리즘을 추천해 주는 MOD 서비스
- Rapidminer 분석 도구인 Operator 활용 시 다수의 Rapidminer User의 사용 이력 정보를 제공
- 개별 Operator들에 대한 정의 및 기능에 대한 Help 메뉴와 활용 예제 제공

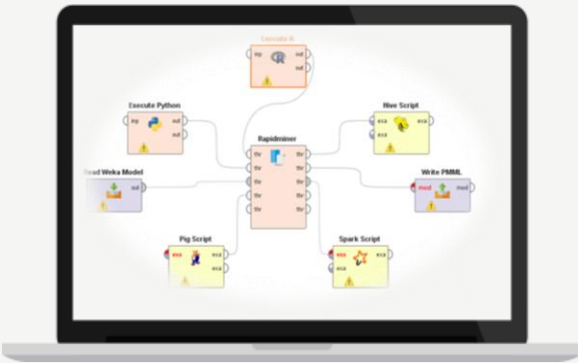


모델링을 위한 다양한 분석 Algorithm 제공

- 기본적인 통계함수 뿐만 아니라 Machine Learning 알고리즘도 제공
- 모델의 최적화를 위한 Optimization Operator
모델에 대한 검증을 위한 Validation Operator
모델에 대한 가중치 및 가시화



유연한 확장성



- SAS, SPSS, DB, URL 등 30여 종류의 데이터 타입을 읽어 들일 수 있는 기능 제공
- R or Python에 익숙한 사용자들을 위해 해당 Script를 직접 작성할 수 있는 환경 제공
- RapidMiner에서 생성된 Model을 타 시스템에 사용할 수 있도록 PMML 저장 기능 제공



RapidMiner Server

Collaboration & Deployment software



주요 기능

- One Click 게시
- Server-based Repository
- 분석 결과 및 프로세스의 스케줄링
- LDAP 유저 권한 설정
- 타 시스템과의 Easy 연동
- Web-service APIs
- 모델 관리

강력한 협업

RapidMiner Server는 RapidMiner Studio를 통해 만들고 저장된 프로세스 뿐만 아니라 분석된 결과물을 쉽게 게시, 공유, 수정 및 관리할 수 있어 팀원간 혹은 팀들간의 강력한 협업을 할 수 있는 환경을 제공합니다.

데이터 분석의 관리

RapidMiner Server는 모델 공유 뿐만 아니라 효과적인 모델 version 관리 및 접근 권한과 같은 분석 관리에 필요한 환경을 제공합니다.

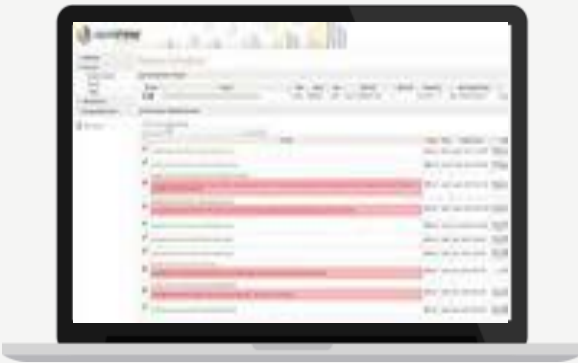
RapidMiner server의 connectors, BI integration 및 Web-services API 등을 통해 분석 결과들을 사용자 비즈니스에 쉽게 통합할 수 있도록 지원합니다.

팀원간의 협업

- 템플릿 과 프로세스를 통해 팀원들간의 관련 지식 공유
- 강력한 유저 관리
- 중앙 Respository 를 통한 모델,프로세스,프로젝트 공유



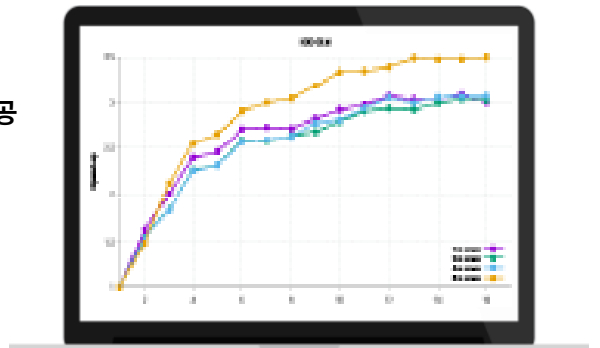
생산적 게시 및 모델 관리



- 결과에 대한 Visualizations 및 비즈니스 Application / Web Services와의 연동
- 정기적 또는 이벤트 기반 모델 실행
- 지속적 모델 업데이트를 통한 정확도 향상

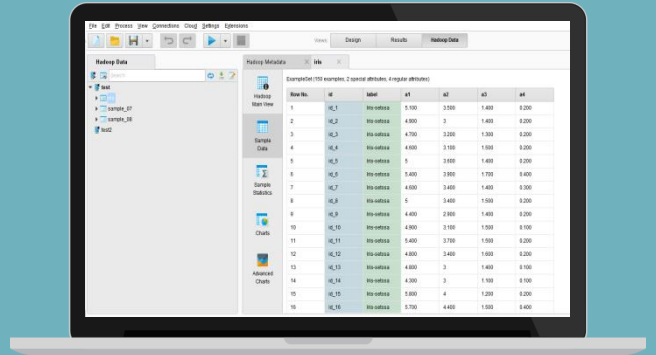
분석 성능의 향상

- Desktop version 보다 높은 사양의 Hardware 지원
- In-Memory 분석 알고리즘과 대용량 데이터 분석 환경 제공
- RapidMiner Studio, Radoop 및 타 데이터 소스들과의 유연한 통합 관리



RapidMiner Radoop

In-Hadoop Visual Workflow Designer



주요 기능

- GUI 프로그래밍 환경
- 많은 데이터 분석 Operator 제공
- Hadoop system 안에서의 자동 실행
- Spark, Hive, Pig 스크립트 사용 가능
- Hadoop 내 분석 모듈 사용으로
대용량 데이터 이동 최소화

쉬운 빅데이터 분석

RapidMiner Radoop은 RapidMiner Studio의 Visual workflow designer로서 Hadoop 시스템 안에서 보다 쉽게 예측 분석이 가능하도록 해주는 모듈입니다. No code 환경과 모듈화된 operator들을 사용함으로써 추가적인 기술적 어려움 없이 복잡한 Hadoop 시스템의 접근을 가능하게 합니다.

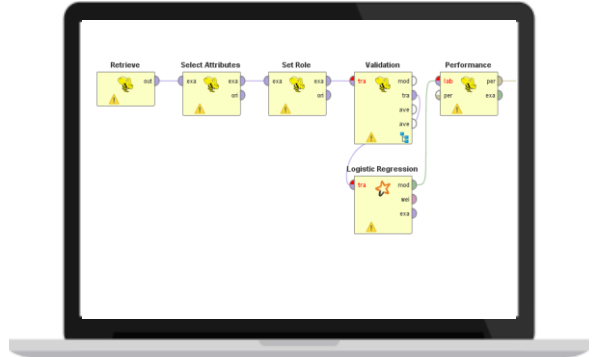
Hadoop 의 가치를 증대

RapidMiner Radoop은 Hadoop의 다양한 분석 모듈을 쉽게 활용 가능하도록 지원하고, GUI기반으로 만들어진 workflow들을 Hive, MapReduce, Spark 등으로 자동 변환 해 주기 때문에 복잡한 분산 처리 기술을 반드시 학습할 필요가 없습니다.

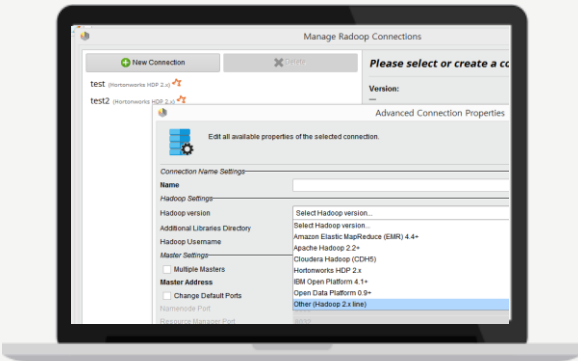
모든 연산은 데이터가 들어있는 Hadoop 클러스터에서 진행이 되기 때문에 대용량 데이터(테라바이트 이상급)를 분석 하는 데에도 효과적입니다.

Hadoop 의 복잡성을 단순화

- Hadoop 분석 업무를 GUI 기반으로 생성
- Process 를 Hadoop 기술 (Hive query, MapReduce & Spark 등) 로 자동 변환
- 모든 클러스터 작동을 관리



효율적 연결 관리



- Wizards 방식의 Big data system 연결
- Major Hadoop 공급 업체 (Cloudera, Hortonworks, Amazon EMR, Apache, IBM Platform) 지원
- 기타 업체들에 대해서도 적절한 Library 설정을 통해 연결 가능

In-Hadoop 과 In-Memory processing의 결합

- 데이터를 메모리와 Hadoop cluster 사이로 교환 가능
- RapidMiner의 수 많은 Operator 사용 가능 (In-Memory)
- 대용량 데이터에 대한 작업을 위해 Hadoop으로 데이터 이동

