

ankus Analyzer G 사용 매뉴얼

onycom
2017.08



Table of Contents



1.

개요

| | |
|----------------------------------|----|
| 1.1 개요----- | p2 |
| 1.2 ankus Analyzer G 개요----- | p3 |
| 1.3 ankus Analyzer G 사용 순서 ----- | p4 |

2.

ankus Analyzer G 화면 설명

| | |
|-------------------------|-----|
| 2.1 로그인----- | p5 |
| 2.2 메인 화면----- | p6 |
| 2.3 관리자 도구----- | p7 |
| 2.4 나의 정보----- | p16 |
| 2.5 진행과정관리 ----- | p19 |
| 2.6 HDFS 브라우저 ----- | p17 |
| 2.7 워크플로우----- | p22 |
| 2.8 모니터링 ----- | p27 |
| 2.9 Hive Query 도구 ----- | p28 |
| 2.10 스케줄러----- | p29 |

시각화

3.

| | |
|---------------|-----|
| 3.1 차트도구----- | p30 |
|---------------|-----|



● ankus Analyzer G 1.0 버전을 기준으로 작성된 사용 매뉴얼

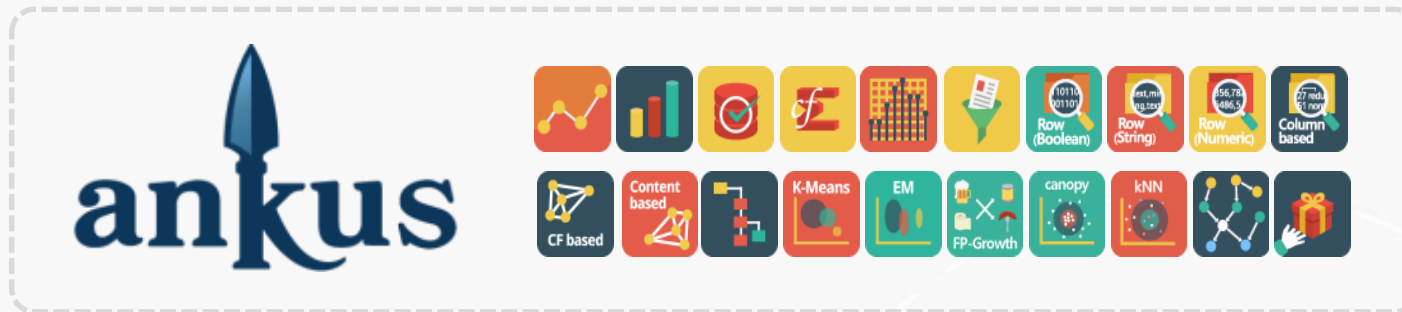
- ※ 본 매뉴얼은 ankus Analyzer G 분석의 사용법 및 각종 화면에 대한 설명을 기술했습니다.
- ※ 본 매뉴얼은 CentOS 6.5 설치된 서버 환경에서 작성/확인되었습니다.
- ※ 필요에 따라 일부 과정을 생략하고 진행할 수 있으며, 이 경우 컴퓨터 설정에 따라 추가적인 작업이 필요할 수 있습니다.
- ※ 본 매뉴얼에서 사용된 **설치 파일 다운로드 및 문의는 ankus@onycom.com**으로 하시기 바랍니다.

본 문서의 저작권은 ankus의 개발/공급을 담당하는 어니컴(주)에 있으며 **임의로 본 문서를 복제/수정하거나 재 배포 할 수 없습니다.**

1.2 ankus Analyzer G 개요



ankus Analyzer G 개요



- 분석 알고리즘과 함수로 구성된 분석 라이브러리(Library)
- Hadoop(대용량 분산 병렬 처리 시스템) 기반 최적화 구현
- 데이터마이닝/머신러닝 기능 지원
- 제품/서비스에 분석 기능 필요 시 활용 적합 (예측 분석, 개인화 추천 지원)

1.3 ankus Analyzer G 사용 순서



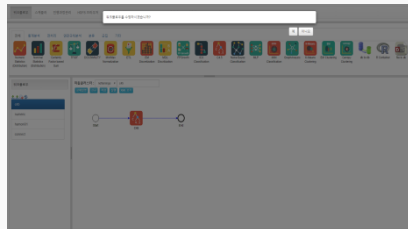
ankus Analyzer G 를 이용해 분석을 수행하는 순서



01

ankus Analyzer G 접속

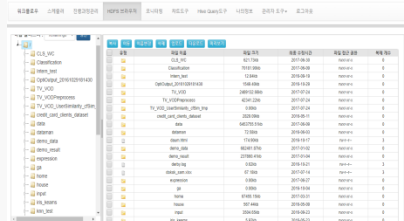
사용자 ID/PW를 이용해 ankus g 시스템에 접속합니다.



05

분석 저장 및 실행

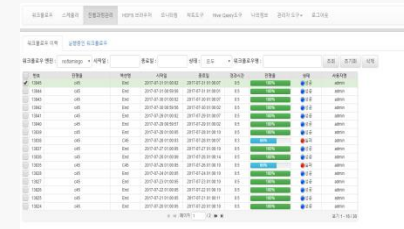
작성된 워크플로우를 저장 후 분석을 수행합니다.



02

데이터 업로드

HDFS 브라우저를 이용해 데이터를 하둡 HDFS에 업로드 합니다.



06

진행과정관리에서 작업 확인

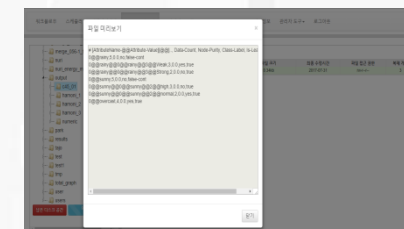
진행과정관리를 통해서 하둡에서의 작업 진행상황을 확인합니다.



03

워크플로우 생성

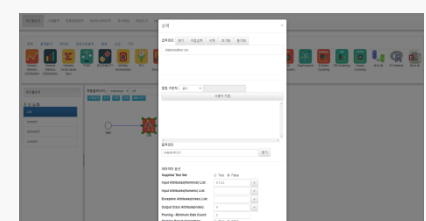
ankus g에서 Drag & Drop 방식으로 분석 알고리즘을 선택하여 워크플로우를 설계합니다



07

분석 결과 확인

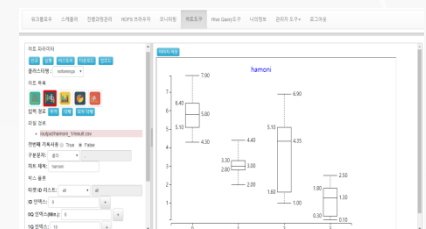
HDFS 브라우저에서 분석 저장 경로로 들어가 분석 결과를 다운로드 합니다.



04

파라미터 설정

데이터 선택, 저장 경로, 변수 선택 등 알고리즘 별로 분석에 필요한 파라미터를 설정합니다.



08

D3.js를 이용한 시각화

차트도구에서 분석 결과에 따라 원하는 차트를 선택하여 D3.js 기반의 시각화를 수행합니다.



- 웹 페이지를 통해 확인할 수 있는 ankus Analyzer G 로그인 화면

- 관리자는 사용자에게 로그인 권한을 부여할 수 있고, 사용자는 ID/Password를 통해 로그인 한다.

ankus Big Data Analytics Platform

ankus

아이디

비밀번호

회원가입 로그인 아이디/비밀번호 찾기

회원가입

아이디

이메일

비밀번호

비밀번호 확인

계정신청 절차

신청양식작성 > 승인요청(메일작성) > 관리자승인 > 신청완료

확인 취소

아이디/비밀번호 찾기

아이디

이메일

확인 취소

Sign up

Find your ID / pw



● 로그인 후 나타나는 ankus Analyzer G 메인 화면

- 로그인 후 상단에 전체 메뉴가 나타난다.

The screenshot shows the main interface of ankus Analyzer G. The top navigation bar includes the following items: 워크플로우, 스케줄러, 진행과정관리, HDFS 브라우저, 모니터링, 차트도구, Hive Query 도구, 나의정보, 관리자 도구, and 로그아웃. Below this, a row of icons represents various analysis tools: Numeric Statistics (Distribution), Nominal Statistics (Distribution), Certainty Factor based Sum, TFIDF, DOCSIMILITY, Min/Max Normalization, ETL, EM Discretization, MDL Discretization, FP-Growth, ID3 Classification, C4.5 Classification, and Naive Bayes Classification. A dropdown menu for '관리자 도구' (Admin Tools) is open, showing options like 사용자관리, 하둡관리, 서버관리, 알고리즘관리, Hive Meta관리, and 언어설정. The main content area shows a workflow editor with a list of tasks on the left (c45, numeric, hamoni01, connect) and a central workspace with a '하둡클러스터' dropdown set to 'noflamingo' and buttons for '상세설명', '실행', '저장', and '시작'. A 'Start' button is also visible. The bottom right corner has an 'End' button.

워크플로우 스케줄러 진행과정관리 HDFS 브라우저 모니터링 차트도구 Hive Query 도구 나의정보 관리자 도구 로그아웃

전체 통계분석 전처리 연관규칙분석 분류 군집 기타

Numeric Statistics (Distribution) Nominal Statistics (Distribution) Certainty Factor based Sum TFIDF DOCSIMILITY Min/Max Normalization ETL EM Discretization MDL Discretization FP-Growth ID3 Classification C4.5 Classification Naive Bayes Classification

로그아웃: ankus 로그아웃

관리자 도구: 관리자 메뉴 드롭다운

나의 정보: 내 계정 관리

Hive Query 도구: SQL on Hadoop 사용

차트도구: 시각화

모니터링: 시스템 상태 확인

HDFS 브라우저: 파일 시스템

진행과정관리: 작업 관리 및 모니터링

스케줄러: 배치 작업 관리

워크플로우: 워크플로우 설계 및 분석수행

하둡클러스터: noflamingo

상세설명 실행 저장 시작

Start

End

사용자관리 하둡관리 서버관리 알고리즘관리 Hive Meta관리 언어설정

2.3 관리자 도구



● 사용자 관리 - 관리자 설정

- 사용자 관리는 'ROLE_ADMIN' 계정을 가진 관리자만 가능하다.
- '활성화' 컬럼에서는 각 사용자의 계정 활성화 여부를 결정할 수 있다.
- '권한' 컬럼에서는 각 사용자에게 부여될 권한을 정할 수 있다.

관리자 도구 > 로그인

사용자관리
하둡관리
서버관리
알고리즘관리
Hive Meta관리
언어설정

사용자명: 이메일: 활성화: 전체 ▼ 권한: 전체 ▼ 생성일시: 조회

'true/false'로 계정 활성화 여부 결정
'USER/ADMIN' 중에서 권한 부여

| 사용자명 | 이메일 | 활성화 | 권한 | 생성일시 | 생성일시 |
|-----------|---------------------|------|------------|--------------------|--------------------|
| rockplace | a@a.com | true | ROLE_ADMIN | 2017-7-27 10:54:45 | 2017-7-13 14:54:36 |
| admin | dodzjtm@dodzjtm.com | true | ROLE_ADMIN | 2017-7-31 9:47:45 | 2015-4-20 9:52:55 |

조건 검색에 따른 조회 버튼

Grid 더블 클릭 후 상세화면 조회

사용자 정보 수정

사용자명: rockplace

이메일: a@a.com

활성화: true ▼

권한: ROLE_ADMIN ▼

이메일 / 활성화 / 권한 수정

변경된 Grid 수정

선택된 Grid 삭제

삭제 수정 취소

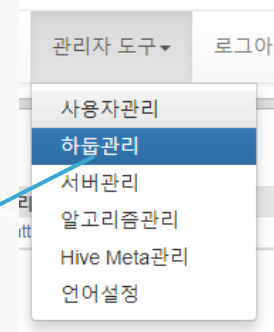
하둡 관리- 관리자 설정 (1/2)

- 관리자 도구 > 하둡 관리 경로로 들어간다.
- 추가 버튼을 누른 후 ankus Analyzer G 와 연동할 하둡 클러스터 정보를 입력한다.
- 등록을 완료하면 하둡 관리 페이지 링크와 함께 목록에 나타난다.

추가

갱신

| 클러스터명 | 네임노드 | 리소스 매니저 | 네임노드 콘솔 링크 | 리소스 매니저 콘솔 링크 |
|------------|---------------------------|--------------------|---|---|
| noflamingo | hdfs://192.168.10.40:9000 | 192.168.10.40:9001 | http://192.168.10.40:50070 | http://192.168.10.40:8055 |



하둡 클러스터 추가

클러스터명

네임노드 & 리소스 매니저

파일 시스템 스키마

네임노드

리소스 매니저

하둡 웹 콘솔

네임노드 콘솔

리소스 매니저 콘솔

cluster 이름(사용자 지정)

파일 시스템 종류(기본 HDFS)

네임노드 서버주소 및 포트(하둡 'fs.default.name' 설정 정보)

잡트래커 서버주소 및 포트(하둡 'mapred.job.tracker' 설정 정보)

DFS 웹 서비스 주소 및 포트

잡트래커 웹 서비스 주소 및 포트

● 하둡 관리 - 관리자 설정 (2/2)

- 등록된 하둡 클러스터가 목록에 나타난다.
- 하둡 관리 페이지 링크를 통해 등록한 하둡의 상태를 확인할 수 있다.

☆ 등록된 Hadoop cluster 목록

Hadoop 관리
페이지 링크

| 추가 | 갱신 | | | |
|------------|---------------------------|--------------------|---|---|
| 클러스터명 | 네임노드 | 리소스 매니저 | 네임노드 콘솔 링크 | 리소스 매니저 콘솔 링크 |
| noflamingo | hdfs://192.168.10.40:9000 | 192.168.10.40:9001 | http://192.168.10.40:50070 | http://192.168.10.40:8055 |



● 서버 관리 - 관리자 설정 (1/2)

- 관리자 도구 > 서버 관리 경로로 들어간다.
- ankus Analyzer G 웹 페이지와 연동할 서버 엔진의 정보를 입력한다.

관리자 도구

로그인

사용자관리

하둡관리

서버관리

알고리즘관리

Hive Meta관리

언어설정

추가

갱신

| 서버명 | 상태 | 하둡 클러스터 | HIVE 서버 | 서버 주소 | 스케줄러 | IP 주소 | 실행 상태 | 접근 권한 |
|------------|---------|------------|---------|---|-----------|---------------|-------|-------|
| noflamingo | RUNNING | noflamingo | | http://192.168.10.40:8080/remote/engine | scheduler | 192.168.10.40 | 0 | |

환경변수

시스템속성

트리거

실행중인 워크플로우

서버 추가

엔진명

엔진

하둡 클러스터

공용

사용자

추가

삭제

취소

Engine 이름(사용자 지정)

Workflow Engine 서버 주소 및 웹 서비스 포트

전 단계에서 등록한 Hadoop cluster 선택

엔진 사용자 권한 체크

특정 사용자만 엔진 사용 권한 부여

● 서버 관리- 관리자 설정 (2/2)

- 엔진 등록이 완료되면 목록에 나타나게 되고, '상태'를 통해 작동 상황을 확인한다.
- 아래 엔진 정보 Tab을 통해 상세 내용을 확인할 수 있다.

★ 등록 된 Engine 목록

추가
갱신

Engine 정보 (Environments, System Properties, Triggers, Running Jobs)

| 서버명 | 상태 | 하둡 클러스터 | HIVE 서버 | 서버 주소 | 스케줄러 | IP 주소 | 실행 상태 | 접근 권한 |
|------------|---------|------------|---------|-------|-----------|-------|-------|-------|
| noflamingo | RUNNING | noflamingo | | | scheduler | | 0 | |

환경변수
시스템속성
트리거
실행중인 워크플로우

| 변수 | 값 |
|------------------------------|-------------------------------------|
| TERM | xterm-256color |
| HADOOP_COMMON_LIB_NATIVE_DIR | /home/ankus/hadoop-2.6.0/lib/native |
| JAVA_HOME | /home/ankus/jdk1.7.0_67 |
| SHLV | 2 |
| HIVE_HOME | /home/ankus/hive |
| XFILESEARCHPATH | /usr/dt/app-defaults/%L/Dt |
| SSH_CLIENT | 192.168.10.106 51444 22 |



알고리즘 관리 - 관리자 설정

- 관리자 도구 > 알고리즘 관리 경로로 들어간다.
- ankus Analyzer G Core를 업로드 하여 관리할 수 있다.
- ankus Analyzer G는 알고리즘은 포함되어 있지 않습니다. Core 파일을 다운로드 받아 업로드 해야 합니다.
- ankus 알고리즘 : <http://www.openankus.com/upload/modules/ankus-core2-onycom-alpha1-1.1.0.jar>
- ankus 차트도구 : <http://www.openankus.com/upload/modules/ankus-visualization-1.0.1.jar>

관리자 도구 ▶ 로그인

사용자관리
하둡관리
서버관리
알고리즘관리
Hive Meta관리
언어설정

| 파일명 | 파일명 | 파일명 |
|---|-------------------------|---|
| <input type="checkbox"/> /tmp/cache/ankus-core2-onycom-alpha1-1.1.0.jar | onycom_Analyzer_package | Numeric Statistics (Distribution), Nominal Statistics (Distribution), Certainty F |
| <input type="checkbox"/> /tmp/cache/db2db.jar | onycom_db2db | db to db |
| <input type="checkbox"/> /tmp/cache/Rcontainer.jar | onycom_RContainer | R Container |
| <input type="checkbox"/> /tmp/cache/file2db.jar | onycom_file2db | file to db |

알고리즘 업로드

업로드 중지 취소 초기화

찾기 Drag & Drop Files

로컬에 위치한 core 파일 업로드 기능

● Hive Meta 관리- 관리자 권한 (1/3)

- 관리자 도구 > Hive Meta 관리 경로로 들어간다.
- ankus Analyzer G 는 사용자가 DataWarehouse(DW)를 쉽게 구축할 수 있도록 Apache Tajo를 연동하여 제공한다.
- Tajo는 대용량 분산 DW 시스템으로서, SQL-on-Hadoop 기능을 제공한다.
- 사용자는 엑셀로 다루기 힘든 큰 데이터를 변환/별도의 Load 없이 바로 SQL로 처리할 수 있다.



위크플로우 스케줄러 진행과정관리 HDFS 브라우저 모니터링 차트도구 Hive Query도구 나의정보 관리자 도구 ▼ 로그아웃

하둡 클러스터 : noflamingo 경신

- CLS_WC
- Classification
- Intern_test
- OptiOutput_20161029181430
- TV_VOD
- TV_VODPreprocess
- TV_VOD_UserSimilarity_cfSim_tmp
- credit_card_clients_dataset
- _getNominalKey 0.00kb
- benchmark 0.00kb
- big_100000_iris.csv 458984.38kb
- big_10000_iris.csv 45898.44kb
- big_1000_1_iris.csv 236621.09kb
- big_5000_50_iris.csv 597167.97kb
- big_iris.csv 4589843.75kb
- graph 0.32kb
- house 755.55kb
- iris.csv 4.44kb
- iris2.csv 6.55kb

Tajo 포트: 26002 최대행 수: 100 파일 미리보기(HDFS)

데이터베이스: default 삭제 데이터베이스명: clustering_data 생성

테이블: 삭제 구분문자: 콤마 첫줄 컬럼 여부: ☐

테이블 생성

| | 0 | 1 | 2 | 3 | 4 |
|-----|-----|-----|-----|-------------|---|
| 5.1 | 3.5 | 1.4 | 0.2 | Iris-setosa | |
| 4.9 | 3.0 | 1.4 | 0.2 | Iris-setosa | |
| 4.7 | 3.2 | 1.3 | 0.2 | Iris-setosa | |
| 4.6 | 3.1 | 1.5 | 0.2 | Iris-setosa | |
| 5.0 | 3.6 | 1.4 | 0.2 | Iris-setosa | |
| 5.4 | 3.9 | 1.7 | 0.4 | Iris-setosa | |
| 4.6 | 3.4 | 1.4 | 0.3 | Iris-setosa | |
| 5.0 | 3.4 | 1.5 | 0.2 | Iris-setosa | |
| 4.4 | 2.9 | 1.4 | 0.2 | Iris-setosa | |
| 4.9 | 3.1 | 1.5 | 0.1 | Iris-setosa | |
| 5.4 | 3.7 | 1.5 | 0.2 | Iris-setosa | |
| 4.8 | 3.4 | 1.6 | 0.2 | Iris-setosa | |
| 4.8 | 3.0 | 1.4 | 0.1 | Iris-setosa | |

● Hive Meta 관리- 관리자 권한 (2/3)

- DW를 생성하기 위해 필요한 데이터 파일을 불러온다.

하둡 클러스터 : noflamingo 갱신

CLS_WC

Classification

Intern_test

OptiOutput_20161029181430

TV_VOD

TV_VODPreprocess

TV_VOD_UserSimilarity_cfSim_tmp

credit_card_clients_dataset

| | | |
|--|----------------------|--------------|
| | _getNominalKey | 0.00kb |
| | benchmark | 0.00kb |
| | big_100000_iris.csv | 458984.38kb |
| | big_10000_iris.csv | 45898.44kb |
| | big_1000_1_iris.csv | 236621.09kb |
| | big_5000_50_iris.csv | 597167.97kb |
| | big_iris.csv | 4589843.75kb |
| | graph | 0.32kb |
| | house | 755.55kb |
| | iris.csv | 4.44kb |
| | iris2.csv | 6.55kb |

★ Engine cluster 선택 후 HDFS Folder 클릭

★ 대상 파일 선택

● Hive Meta 관리- 관리자 권한 (3/3)

- 생성할 데이터베이스의 META정보를 입력하고 미리보기 기능을 통해 확인한다.



- Tajo Port : Tajo Port 지정
- 최대행 수: Preview 시 출력 라인수
- 파일미리보기(HDFS) : 하단 화면 preview
- 데이터베이스: 생성된 DB 선택
- 데이터베이스명: 새로운 DB 명 작성 후 create 버튼으로 생성
- 테이블: 생성된 Table 선택
- 구분문자: file 구분자 선택
- 첫줄 컬럼 여부: 첫번째 row 사용여부
- 테이블 생성: table 생성



- 테이블 명: 테이블 명 입력
- 속성: 각 행의 이름 및 속성을 입력

Tajo 포트: 최대행 수: 파일 미리보기(HDFS)

데이터베이스: 삭제 데이터베이스명: 생성

테이블: 삭제 구분문자: 첫줄 컬럼 여부: ☐

| 0 | 1 | 2 | 3 | 4 |
|-----|-----|-----|-----|-------------|
| 5.1 | 3.5 | 1.4 | 0.2 | Iris-setosa |
| 4.9 | 3.0 | 1.4 | 0.2 | Iris-setosa |
| 4.7 | 3.2 | 1.3 | 0.2 | Iris-setosa |
| 4.6 | 3.1 | 1.5 | 0.2 | Iris-setosa |
| 5.0 | 3.6 | 1.4 | 0.2 | Iris-setosa |
| 5.4 | 3.9 | 1.7 | 0.4 | Iris-setosa |
| 4.6 | 3.4 | 1.4 | 0.3 | Iris-setosa |
| 5.0 | 3.4 | 1.5 | 0.2 | Iris-setosa |
| 4.4 | 2.9 | 1.4 | 0.2 | Iris-setosa |
| 4.9 | 3.1 | 1.5 | 0.1 | Iris-setosa |
| 5.4 | 3.7 | 1.5 | 0.2 | Iris-setosa |
| 4.8 | 3.4 | 1.6 | 0.2 | Iris-setosa |
| 4.8 | 3.0 | 1.4 | 0.1 | Iris-setosa |

테이블 생성

테이블명:

속성0:

속성1:

속성2:

속성3:

속성4:

2.4 나의 정보



● 나의 정보- 사용자 설정

- 로그인 한 계정에 대한 정보를 변경 가능하다.(패스워드 재설정, e메일 주소 변경 가능)

The image displays the '나의 정보' (My Information) settings page in the ankus system. The page has a top navigation bar with 'Query도구', '나의정보', and '관리자' tabs. The '나의정보' tab is selected. The main content area is titled '나의 정보' and contains several input fields: '아이디' (ID) with the value 'admin', '현재 비밀번호' (Current Password) with a masked input, '새 비밀번호' (New Password), '새 비밀번호 확인' (Confirm New Password), and '이메일' (Email) with the value 'dodzjtm@dodzjtm.com'. There are '확인' (Confirm), '적용' (Apply), and '취소' (Cancel) buttons. A red arrow points to the '확인' button with the text '현재 패스워드 입력' (Enter current password). Another red arrow points to the '적용' button with the text '정보 변경 칸 활성화' (Activate information change field). A blue arrow points from the '나의정보' tab to the '나의 정보' form.

Query도구 나의정보 관리자

나의 정보

아이디 admin

현재 비밀번호 확인

새 비밀번호

새 비밀번호 확인

이메일 dodzjtm@dodzjtm.com

적용 취소

현재 패스워드 입력

정보 변경 칸 활성화



● 분석 작업 상태 모니터링

- 먼저, 워크플로우 엔진을 선택하고 해당 엔진에서 수행한 분석 작업 목록을 확인한다.
- 상태를 통해 작업 성공/실패 여부를 확인할 수 있다.

줄러 **진행과정관리** HDFS 브

워크플로우 이력 실행중인 워크플로우

워크플로우 엔진 : **noflamingo** 시작일 : 종료일 : 상태 : **모두** 워크플로우명 : 조회 초기화 삭제

| 번호 | 진행율 | 액션명 | 시작일 | 종료일 | 경과시간 | 진행율 | 상태 | 사용자명 |
|-------|-----|-----|---------------------|---------------------|------|------|----|-------|
| 13845 | c45 | End | 2017-07-31 01:00:02 | 2017-07-31 01:00:07 | 0.5 | 100% | 성공 | admin |
| 13844 | c45 | End | 2017-07-31 00:59:56 | 2017-07-31 01:00:01 | 0.5 | 100% | 성공 | admin |
| 13843 | c45 | End | 2017-07-30 01:00:02 | 2017-07-30 01:00:07 | 0.5 | 100% | 성공 | admin |
| 13842 | c45 | End | 2017-07-30 00:59:56 | 2017-07-30 01:00:02 | 0.5 | 100% | 성공 | admin |
| 13841 | c45 | End | 2017-07-29 01:00:02 | 2017-07-29 01:00:07 | 0.5 | 100% | 성공 | admin |
| 13840 | c45 | End | 2017-07-29 00:59:57 | 2017-07-29 01:00:02 | 0.5 | 100% | 성공 | admin |
| 13839 | c45 | End | 2017-07-28 01:00:05 | 2017-07-28 01:00:10 | 0.5 | 100% | 성공 | admin |
| 13838 | c45 | C45 | 2017-07-28 01:00:03 | 2017-07-28 01:00:07 | 0.3 | 66% | 실패 | admin |
| 13837 | c45 | End | 2017-07-27 01:00:05 | 2017-07-27 01:00:10 | 0.5 | 100% | 성공 | admin |
| 13836 | c45 | End | 2017-07-26 01:00:09 | 2017-07-26 01:00:14 | 0.5 | 100% | 성공 | admin |
| 13835 | c45 | C45 | 2017-07-26 01:00:05 | 2017-07-26 01:00:10 | 0.5 | 66% | 실패 | admin |
| 13828 | c45 | End | 2017-07-24 01:00:05 | 2017-07-24 01:00:10 | 0.5 | 100% | 성공 | admin |
| 13827 | c45 | End | 2017-07-23 01:00:05 | 2017-07-23 01:00:10 | 0.5 | 100% | 성공 | admin |
| 13826 | c45 | End | 2017-07-22 01:00:05 | 2017-07-22 01:00:10 | 0.5 | 100% | 성공 | admin |
| 13825 | c45 | End | 2017-07-21 01:00:05 | 2017-07-21 01:00:11 | 0.5 | 100% | 성공 | admin |
| 13824 | c45 | End | 2017-07-20 01:00:05 | 2017-07-20 01:00:10 | 0.5 | 100% | 성공 | admin |

13838, 13835, 13836, 13837, 13838, 13839, 13840, 13841, 13842, 13843, 13844, 13845

작업 목록

1 / 2 보기 1 - 16 / 30



● 작업의 수행 내용 및 Workflow XML 확인

- 작업을 더블 클릭해서 창을 열면 상세한 작업 수행 내용이 확인 가능하다.
- 아래 Tab에서는 워크 플로우 XML과 에러 로그를 볼 수 있다.

워크플로우 이력

실행중인 워크플로우

이크플로우 예지 .

이자인 .

종료일 :

상태 : 모두

워크플로우명 :

조회

초기화

삭제

c45 - WF_20170623_928516800

작업

액션

작업 정보

작업 ID: -1176892425

시작일: 2017-07-31 01:00:02

워크플로우 ID: WF_20170623_928516800

종료일: 2017-07-31 01:00:07

워크플로우명: c45

경과시간: 0.5

상태: 성공

진행율:

워크플로우 XML

에러 로그

```

1 <!--?xml version="1.0" encoding="UTF-8" standalone="yes"?-->
2 <workflow workflowname="c45"
3   xmlns="http://www.openankus.org/schema/workflow">
4   <description></description>
5   <start to="06_7295_8" name="06_7295_2" description="Start">
6   <action to="06_7295_5" name="06_7295_8" description="C45">
7     <mapreduce>
8       <jar>org.ankus:ankus-core2-onycom-alpha1:1.1.0</jar>
9       <classname>C45</classname>
10      <command></command>
11      <variable>
12        <value>-nominalIndexList</value>
13      </variable>
14      <variable>
15        <value>0.1.2.3</value>

```

| 시작일 | 종료일 | 경과시간 | 진행율 | 상태 | 사용자명 |
|---------------------|---------------------|------|------|----|-------|
| 2017-07-31 01:00:02 | 2017-07-31 01:00:07 | 0.5 | 100% | 성공 | admin |
| 2017-07-31 00:59:56 | 2017-07-31 01:00:01 | 0.5 | 100% | 성공 | admin |
| 2017-07-30 01:00:02 | 2017-07-30 01:00:07 | 0.5 | 100% | 성공 | admin |
| 2017-07-30 00:59:56 | 2017-07-30 01:00:02 | 0.5 | 100% | 성공 | admin |
| 2017-07-29 01:00:02 | 2017-07-29 01:00:07 | 0.5 | 100% | 성공 | admin |
| 2017-07-29 00:59:57 | 2017-07-29 01:00:02 | 0.5 | 100% | 성공 | admin |
| 2017-07-28 01:00:05 | 2017-07-28 01:00:10 | 0.5 | 100% | 성공 | admin |
| 2017-07-28 01:00:03 | 2017-07-28 01:00:07 | 0.3 | 66% | 실패 | admin |
| 2017-07-27 01:00:05 | 2017-07-27 01:00:10 | 0.5 | 100% | 성공 | admin |
| 2017-07-26 01:00:09 | 2017-07-26 01:00:14 | 0.5 | 100% | 성공 | admin |
| 2017-07-26 01:00:05 | 2017-07-26 01:00:10 | 0.5 | 66% | 실패 | admin |
| 2017-07-24 01:00:05 | 2017-07-24 01:00:10 | 0.5 | 100% | 성공 | admin |
| 2017-07-23 01:00:05 | 2017-07-23 01:00:10 | 0.5 | 100% | 성공 | admin |
| 2017-07-22 01:00:05 | 2017-07-22 01:00:10 | 0.5 | 100% | 성공 | admin |
| 2017-07-21 01:00:05 | 2017-07-21 01:00:11 | 0.5 | 100% | 성공 | admin |
| 2017-07-20 01:00:05 | 2017-07-20 01:00:10 | 0.5 | 100% | 성공 | admin |

워크플로우 XML

에러 로그

1 <!--?xml version="1.0" encoding="UTF-8" standalone="yes"?-->

2 <workflow workflowname="c45"

3 xmlns="http://www.openankus.org/schema/workflow">

4 <description></description>

5 <start to="06_7295_8" name="06_7295_2" description="Start">

6 <action to="06_7295_5" name="06_7295_8" description="C45">

7 <mapreduce>

8 <jar>org.ankus:ankus-core2-onycom-alpha1:1.1.0</jar>

9 <classname>C45</classname>

10 <command></command>

11 <variable>

12 <value>-nominalIndexList</value>

13 </variable>

14 <variable>

15 <value>0.1.2.3</value>

click

워크플로우 XML



● 실패한 작업의 에러 로그 확인

- 상태가 실패인 작업의 에러 로그를 확인할 수 있다.
- 작업이 실패했을 경우, 에러 로그를 통해 실패 원인을 파악한다.

워크플로우 이력 실행중인 워크플로우

워크플로우 엔진 : noflamingo 시작일 : 종료일 : 상태 : 모두 워크플로우명 : 조회 초기화 삭제

c45-WF_20170623_928516800

작업 액션

작업 정보

종료일 : 2017-07-28 01:00:07

워크플로우명 : c45
경과시간 : 0:3

상태 : 실패
진행율 :

최종 액션 : C45
사용자명 : admin

| 시작일 | 종료일 | 경과시간 | 진행율 | 상태 | 사용자명 |
|---------------------|---------------------|------|------|----|-------|
| 2017-07-31 01:00:02 | 2017-07-31 01:00:07 | 0:5 | 100% | 성공 | admin |
| 2017-07-31 00:59:56 | 2017-07-31 01:00:01 | 0:5 | 100% | 성공 | admin |
| 2017-07-30 01:00:02 | 2017-07-30 01:00:07 | 0:5 | 100% | 성공 | admin |
| 2017-07-30 00:59:56 | 2017-07-30 01:00:02 | 0:5 | 100% | 성공 | admin |
| 2017-07-29 01:00:02 | 2017-07-29 01:00:07 | 0:5 | 100% | 성공 | admin |
| 2017-07-29 00:59:57 | 2017-07-29 01:00:02 | 0:5 | 100% | 성공 | admin |
| 2017-07-28 01:00:05 | 2017-07-28 01:00:10 | 0:5 | 100% | 성공 | admin |
| 2017-07-28 01:00:03 | 2017-07-28 01:00:07 | 0:3 | 66% | 실패 | admin |
| 2017-07-27 01:00:05 | 2017-07-27 01:00:10 | 0:5 | 100% | 성공 | admin |
| 2017-07-26 01:00:09 | 2017-07-26 01:00:14 | 0:5 | 100% | 성공 | admin |
| 2017-07-26 01:00:05 | 2017-07-26 01:00:10 | 0:5 | 66% | 실패 | admin |
| 2017-07-24 01:00:05 | 2017-07-24 01:00:10 | 0:5 | 100% | 성공 | admin |
| 2017-07-23 01:00:05 | 2017-07-23 01:00:10 | 0:5 | 100% | 성공 | admin |
| 2017-07-22 01:00:05 | 2017-07-22 01:00:10 | 0:5 | 100% | 성공 | admin |
| 2017-07-21 01:00:05 | 2017-07-21 01:00:11 | 0:5 | 100% | 성공 | admin |
| 2017-07-20 01:00:05 | 2017-07-20 01:00:10 | 0:5 | 100% | 성공 | admin |

워크플로우 XML 에러 로그

Cannot execute '[C45]' of action

```

1 org.ankus.core.exception.WorkflowException: Cannot execute '[C45]' of action
2   at
3   org.ankus.engine.handler.DefaultHandler.execute(DefaultHandler.java:121)
4   at
5   org.ankus.engine.scheduler.DefaultWorkflowJob.executeInternal(DefaultWorkflowJob.java:109)
6   at
7   org.ankus.engine.scheduler.DefaultQuartzJob.execute(DefaultQuartzJob.java:45)
8   at org.quartz.core.JobRunShell.run(JobRunShell.java:202)
9   at
10  org.quartz.simpl.SimpleThreadPool$WorkerThread.run(SimpleThreadPool.java:573)
11  Caused by: org.ankus.core.exception.WorkflowException: MapReduce Job 'C45' is
12  failed. Cause: Process를 실행할 수 없습니다. 프로세스의 종료 코드는 '255'입니
13  다. 로그 : at
14  org.apache.hadoop.hdfs.server.namenode.FSNamesystem.getBlockLocationsUpdateI
  
```

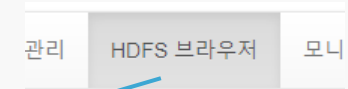
에러 로그

click



● 하둡 파일 시스템(HDFS) 관리 화면

- 먼저 하둡 클러스터에서 등록한 Engine을 선택하여 폴더 목록을 띄운다.
- 파일 관리는 관리자(admin) 또는 사용자(user)의 권한에 따라 기능이 상이하다.
- /user 아래에 사용자 계정명으로 된 경로를 홈 디렉토리(Home directory)로 사용한다.



Hadoop cluster 선택

※ 사용자의 파일 관리 기능 권한

- Copy (홈 디렉토리에서만 복사가능)
- Move (홈 디렉토리 내에서만 이동가능)
- Rename (홈 디렉토리 내에서만 변경가능)
- Delete (홈 디렉토리 내에서만 삭제가능)
- Upload (홈 디렉토리 내로 업로드 가능)

워크플로우 스케줄러 진행과정관리 HDFS 브라우저 모니터링 차트도구 Hive Query도구 나의정보 관리자 도구

하둡 클러스터: noflamingo

복사 이동 이름변경 삭제 업로드 다운로드 미리보기

클릭

HDFS 폴더목록

HDFS 사용량

HDFS 저장 용량 확인

디렉토리 정보

기본 정보

이름: barchart.txt
경로: /input/barchart.txt
유형: 파일 디렉토리
총 크기: 0.08KB
수정일: 2016-04-26 14:30:18

접근 권한

소유자: 읽기 쓰기 실행
그룹: 읽기 쓰기 실행
기타: 읽기 쓰기 실행

공간 사용 정보

읽기: 134217728Bit
디스크 할당량: -1
디스크 사용량(복제포함): 255 B
복제 개수: 3
디렉토리 개수: 0
디렉토리 할당량: -1
파일 개수: 1

닫기

● 데이터 파일 업로드

- 먼저, 사용할 데이터를 HDFS로 옮기기 위해 Upload 창을 연다.
- HDFS로 업로드 할 파일 목록을 불러온다.
- 아래 최대 업로드 사이즈를 확인하여 업로드 할 파일을 선택한다.
- 마지막으로 업로드 결과(진행상태 100%)를 확인한 후 창을 닫는다.



위크플로우 스케줄러 진행과정관리 HDFS 브라우저 모니터링 차트도구 Hive Query도구 나의정보 관리자 도구▼ 로그아웃

하둡 클러스터 : notflamingo [갱신]

CLS_WC
Classification
Intern_test
OptiOutput_20161029181430
TV_VOD
TV_VODPreprocess
TV_VOD_UserSimilarity_cfSim
credit_card_clients_dataset
data
dataman
demo_data
demo_result
expression
ga
home
house
input
Iris_keams

복사 이동 이름변경 삭제 업로드 다운로드 미리보기

파일 업로드

업로드 취소 초기화

찾기 Drag & Drop Files

2) cfsum.txt (0.07 KB)
Cancel

1) cf_sample.txt (0.13 KB)
Cancel

닫기

| 최종 수정시간 | 파일 접근 권한 | 복제 개수 |
|------------|-----------|-------|
| 2016-08-11 | rw-r--r-- | 3 |
| 2016-04-26 | rw-r--r-- | 3 |
| 2016-04-29 | rw-r--r-- | 3 |
| 2016-04-29 | rw-r--r-- | 3 |
| 2016-04-29 | rw-r--r-- | 3 |
| 2016-04-29 | rw-r--r-- | 3 |
| 2016-04-29 | rw-r--r-- | 3 |
| 2016-04-27 | rw-r--r-- | 3 |
| 2016-04-29 | rw-r--r-- | 3 |
| 2016-07-18 | rw-r--r-- | 3 |
| 2016-04-29 | rw-r--r-- | 3 |
| 2016-04-29 | rw-r--r-- | 3 |
| 2016-04-29 | rw-r--r-- | 3 |
| 2016-04-29 | rw-r--r-- | 3 |

● 워크 플로우 화면

- ankus G 분석 라이브러리에서 원하는 알고리즘을 선택해 분석 워크 플로우를 설계한다.
- 설계한 워크 플로우를 저장하면 왼쪽 워크 플로우 목록에 나타난다.
- 필요 시 해당 워크 플로우를 불러와 다시 사용할 수 있다.

워크플로우 스케일러 진행과정관리 HDFS 브라우저 모니터링 차트도구 Hive Query 도구 나의정보 관리자 도구 로그아웃

통계분석 연관규칙분석 분류 군집 기타

Numeric Statistics (Distribution) Nominal Statistics (Distribution) Certainty Factor based Sum TFIDF DOCSIMILITY Min/Max Normalization ETL EM Discretization MDL Discretization FP-Growth FPGrowth ID3 Classification C4.5 Naive Bayes Classification MLP kNN Classification Graph

워크플로우

c45 numeric hamoni01 connect

하둡클러스터 : noflamingo c45

상세설명 신규 저장 실행 XML 보기

Start C4.5 End

워크플로우 디자인 공간



● 분석 워크플로우 작성

- 먼저, 하둡 클러스터를 선택한다.
- 알고리즘 아이콘을 끌어다 워크 플로우 디자인 공간에 놓는다.
- Start 노드와 알고리즘 아이콘을 연결하고, 아이콘과 End 노드를 연결하여 흐름을 완성한다.
- 알고리즘 아이콘을 더블 클릭하여 작업 수행을 위한 필수 정보를 입력한다.

The screenshot shows the ankus workflow editor interface. At the top, there's a navigation bar with tabs: 워크플로우, 스케줄러, 진행과정관리, HDFS 브라우저, 모니터링, 차트도구, Hive Query도구, 나의정보, 관리자 도구, and 로그아웃. Below this, there's a category bar with tabs: 전체, 통계분석, 전처리, 연관규칙분석, 분류, 군집, and 기타. The '분류' (Classification) tab is selected, showing various algorithm icons like Numeric Statistics, Nominal Statistics, Certainty Factor based Sum, TFIDF, DOCSIMILTY, Min/Max Normalization, ETL, EM Discretization, MDL Discretization, FP-Growth, ID3 Classification, C4.5, Naive Bayes Classification, MLP, and kNN Classification. The 'C4.5' icon is highlighted with a red circle and a red arrow pointing to it from the text '워크플로우 이름 입력' (Input workflow name). Another red arrow points from the text 'Drag & Drop' to the 'C4.5' icon. A third red arrow points from the text 'click' to the 'C4.5' icon. In the bottom right, a red star icon is next to the text '설정화면 띄우기' (Open settings screen). The main workspace shows a workflow diagram with a 'Start' node, a 'C4.5' algorithm node, and an 'End' node. The 'C4.5' node is circled in red. On the left, there's a sidebar with a list of workflows: c45, numeric, hamoni01, and connect. The 'c45' workflow is selected. Above the workflow list, there's a dropdown menu showing 'noflamingo' and 'c45'. Below the dropdown, there are buttons: 상세설명, 신규, 저장, 실행, and XML 보기.

● 분석 Parameter 설정

- 먼저, 분석할 데이터를 불러온다.
- 분석 결과를 출력할 경로를 설정한다. (반드시 중복되지 않은 새로운 경로여야 한다)
- 컬럼 구분자를 설정한다. (아래 사용자 지정 버튼을 통해 데이터 모양을 확인한다)
- 분석에 사용할 Target 속성 선택 및 해당 알고리즘에 맞는 나머지 설정을 한다.

★ 분석 데이터 불러오기

데이터 구분자 설정
데이터 미리보기

★ 분석결과 출력 경로 설정

상세

입력경로: 직접입력 삭제 초기화 동기화

컬럼 구분자: 콤마

| 0 column | 1 column | 2 column | 3 column |
|----------|----------|----------|----------|
| sunny | hot | high | Weak |
| sunny | hot | high | Strong |
| overcast | hot | high | Weak |
| rainy | mild | high | Weak |
| rainy | cool | normal | Weak |
| rainy | cool | normal | Strong |

출력경로

파라미터 옵션

Supplied Test Set: ☐ True ☒ False

Input Attributes(Nominal) List:

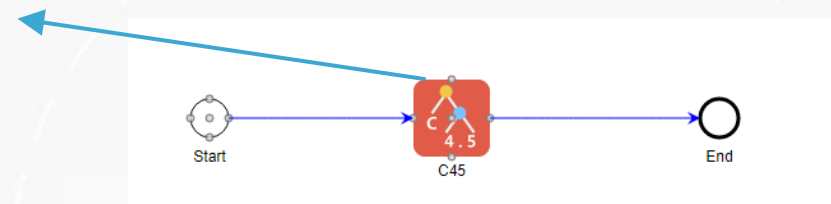
Input Attributes(Numeric) List:

Exception Attributes(Index) List:

Output Class Attribute(Index):

Pruning - Minimum Data Count:

Training Result Generation: ☐ True ☒ False



★ 분석 속성 선택 (타겟 속성/예외 속성)

★ 분석 알고리즘 별로 상이한 추가 설정

2.7 워크 플로우



● 워크 플로우 저장 및 작업 실행

- 파라미터 설정을 완료하고, 작성한 분석 워크 플로우를 저장한다.
- 입력한 워크 플로우 이름이 왼쪽의 워크 플로우 리스트 표시되는지 확인한다.
- 실행 버튼을 눌러 작업을 실행한다.

The screenshot displays the Ankus workflow editor. At the top, a navigation bar includes links for '워크플로우' (Workflows), '스케줄러' (Scheduler), '진행과정관리' (Progress Management), 'HDFS 브라우저' (HDFS Browser), '모니터링' (Monitoring), '차트도구' (Chart Tools), 'Hive Query도구' (Hive Query Tools), '나의정보' (My Info), '관리자 도구' (Admin Tools), and '로그아웃' (Logout). Below this, a toolbar contains various data processing and machine learning icons. The main workspace shows a workflow diagram with a 'Start' node, a 'C4.5' classifier node, and an 'End' node. The 'C4.5' node is highlighted, and its configuration panel on the right shows the '하둡클러스터' (Hadoop Cluster) set to 'hadoop' and 'c45'. The '실행' (Execute) button is visible. On the left, a list of saved workflows includes 'c45', 'numeric', 'hamoni01', and 'connect'. Two blue arrows point from the text instructions to the 'C4.5' node and the '실행' button.

● 작업 진행 및 분석결과 다운로드

- 작업 진행 상황을 진행 과정 관리에서 확인한다. (2.5 진행 과정 관리 참조)
- 작업이 완료되면, HDFS 브라우저를 열어 이전에 설정한 결과 출력 경로로 들어간다.
- 결과 파일을 선택하여 Download 한다.

워크플로우 이력 실행중인 워크플로우

워크플로우 엔진: **noflamingo** 시작일: 종료일: 상태: 모두 워크플로우명: **조회** *새로고침 버튼*

| 번호 | 진행률 | 액션명 | 시작일 | 종료일 | 경과시간 | 진행률 | 상태 | 사용자명 |
|-------|-----|-----|---------------------|---------------------|------|------|----|-------|
| 13845 | c45 | End | 2017-07-31 01:00:02 | 2017-07-31 01:00:07 | 0:5 | 100% | 성공 | admin |
| 13844 | c45 | End | 2017-07-31 00:59:56 | 2017-07-31 01:00:01 | 0:5 | 100% | 성공 | admin |
| 13843 | c45 | End | 2017-07-30 01:00:01 | | | | | |
| 13842 | c45 | End | 2017-07-30 01:00:01 | | | | | |
| 13841 | c45 | End | 2017-07-29 01:00:01 | | | | | |
| 13840 | c45 | End | 2017-07-29 01:00:01 | | | | | |
| 13839 | c45 | End | 2017-07-28 01:00:01 | | | | | |
| 13838 | c45 | End | 2017-07-28 01:00:01 | | | | | |
| 13837 | c45 | End | 2017-07-27 01:00:01 | | | | | |
| 13836 | c45 | End | 2017-07-26 01:00:01 | | | | | |
| 13835 | c45 | End | 2017-07-26 01:00:01 | | | | | |
| 13828 | c45 | End | 2017-07-24 01:00:01 | | | | | |
| 13827 | c45 | End | 2017-07-23 01:00:01 | | | | | |
| 13826 | c45 | End | 2017-07-22 01:00:01 | | | | | |
| 13825 | c45 | End | 2017-07-21 01:00:01 | | | | | |
| 13824 | c45 | End | 2017-07-20 01:00:01 | | | | | |

워크플로우 스케줄러 진행과정관리 HDFS 브라우저 **2** 차트도구 Hive Query도구 나의정보 관리자 도구 로그아웃

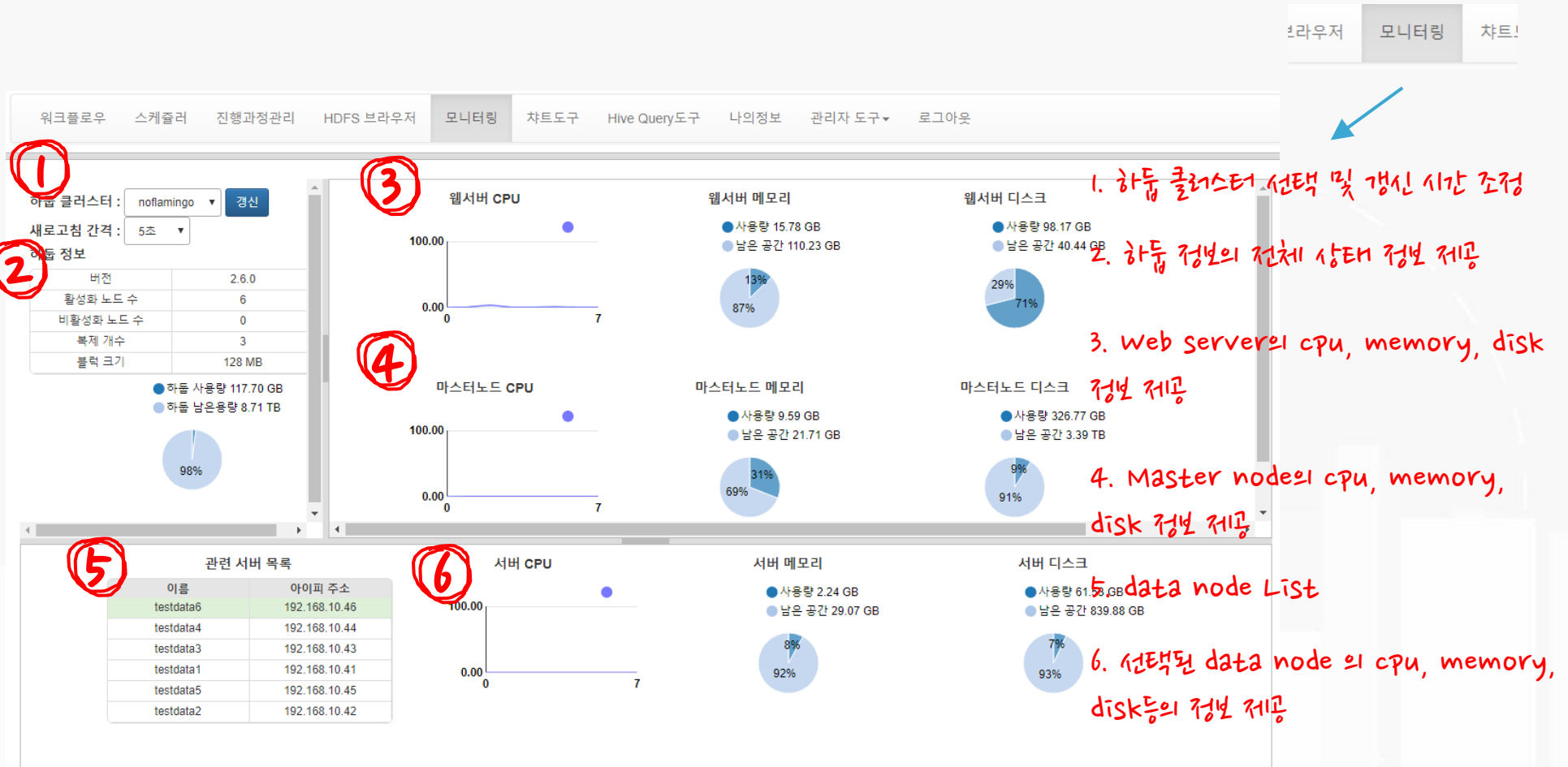
하둡 클러스터 **noflamingo** 경신

복사 이동 이름변경 삭제 업로드 **다운로드** 미리보기

| 유형 | 파일 이름 | 파일 크기 | 최종 수정시간 | 파일 접근 권한 | 복제 개수 |
|----|----------------------------|-----------|------------|-----------|-------|
| | 20160205_115359.jpg | 3316.89kb | 2016-08-11 | rw-r--r-- | 3 |
| | barchart.txt | 0.08kb | 2016-04-26 | rw-r--r-- | 3 |
| | boolean_sample.txt | 0.17kb | 2016-04-29 | rw-r--r-- | 3 |
| | cf_sample.txt | 0.13kb | 2016-04-29 | rw-r--r-- | 3 |
| | csum.txt | 0.07kb | 2016-04-29 | rw-r--r-- | 3 |
| | clustering_data_sample.txt | 0.09kb | 2016-04-29 | rw-r--r-- | 3 |
| | contents_sample.txt | 0.12kb | 2016-04-29 | rw-r--r-- | 3 |
| | credit.txt | 148.13kb | 2016-04-27 | rw-r--r-- | 3 |
| | diabetes.txt | 32.97kb | 2016-04-29 | rw-r--r-- | 3 |
| | iris.txt | 4.59kb | 2016-07-18 | rw-r--r-- | 3 |
| | lenses.txt | 0.85kb | 2016-04-29 | rw-r--r-- | 3 |
| | numeric_sample.txt | 0.17kb | 2016-04-29 | rw-r--r-- | 3 |
| | string_sample.txt | 0.03kb | 2016-04-29 | rw-r--r-- | 3 |
| | weather_tennis.txt | 0.37kb | 2016-04-29 | rw-r--r-- | 3 |

● 모니터링

- 웹 서버 및 Hadoop System 의 상태 정보를 제공한다.



● Hive Query 도구- 사용자 권한

- Hive Query 도구로 들어간다.
- Hive Query 도구에서는 Tajo SQL을 작성하여 적용할 수 있다.

워크플로우 스케줄러 진행과정관리 HDFS 브라우저 모니터링 차트도구 **Hive Query 도구** 나의정보 관리자 도구▼ 로그아웃

하둡 클러스터 : noflamingo ▼ Tajo 포트 : 26002

데이터베이스 : work ▼ 테이블 : weather ▼ 최대행 수 : 100

SELECT * FROM weather LIMIT 100

쿼리 실행

| | a1 | a2 | a3 | a4 | class |
|----|----------|------|--------|--------|-------|
| 1 | sunny | hot | high | Weak | no |
| 2 | sunny | hot | high | Strong | no |
| 3 | overcast | hot | high | Weak | yes |
| 4 | rainy | mild | high | Weak | yes |
| 5 | rainy | cool | normal | Weak | yes |
| 6 | rainy | cool | normal | Strong | no |
| 7 | overcast | cool | normal | Strong | yes |
| 8 | sunny | mild | high | Weak | no |
| 9 | sunny | cool | normal | Weak | yes |
| 10 | rainy | mild | normal | Weak | yes |
| 11 | sunny | mild | normal | Strong | yes |



- 하둡 클러스터 : Engine 선택
- Tajo 포트 : Tajo Port 지정
- 데이터베이스 : 생성된 DB 선택
- 테이블 : 생성된 Table 선택
- 최대행 수: query 결과 출력 라인 수
- 쿼리: list를 보여주는 query command
- 쿼리 실행: query 실행

2.10 스케줄러



● 스케줄러 - 사용자 권한

- 메뉴에 스케줄러를 클릭한다.
- 스케줄 종류(workflow / 로컬)로 구분하며 사용여부로 가동 여부를 결정한다.
- 월간/ 주간/ 일간/ 시간/ 분간 으로 구동 시간을 결정 한다.

워크플로우 스케줄러 진행과정관리 HDFS 브라우저 모니터링 차트도구 Hive Query도구 나의정보 관리자 도구▼ 로그아웃

스케줄 종류: 전체 스케줄명: 워크플로우명: 사용여부: 전체 조회

생성 **삭제**

| 스케줄명 | 스케줄 종류 | 스케줄 | 월 | 주 | 일 | 시 | 분 | 상태 | 실행시간 | 종료시간 | 경과시간 | 사용여부 | 수동실행 |
|--------|----------|---------|---|---|---|---|---|----|-----------------|-----------------|------|------|------|
| 의사결정트리 | workflow | 스케줄러 생성 | | | | 0 | | | 2017-7-31 1:0:8 | 2017-7-31 1:0:8 | 0:0 | Y | |
| 수집기 | 로컬실행 | | | | | 0 | | | | | | N | |

1 / 1 20 보기 1 - 2 / 2

스케줄명: 스케줄 종류: 워크플로우: 월: ☒ 매일 ☐ 1월 ☐ 2월 ☐ 3월 ☐ 4월 ☐ 5월 ☐ 6월 ☐ 7월 ☐ 8월 ☐ 9월 ☐ 10월 ☐ 11월 ☐ 12월 요일: ☒ 매일 ☐ 일 ☐ 월 ☐ 화 ☐ 수 ☐ 목 ☐ 금 ☐ 토 일(1~31): ex) 1,5,10,15,20,25,30 (empty is every) 시(0~23): 1 분(0~59): 0 사용여부: Y

생성 **취소**

선택된 스케줄러를 수동 실행

- 스케줄명: 스케줄 명을 입력
- 스케줄 종류: 워크플로우 / 로컬 선택
- 워크플로우: 스케줄 종류를 워크플로우로 선택시 저장된 워크플로우를 호출
- 월 / 요일 / 일 / 시 / 분 의 수행 주기를 입력 및 선택
- 사용여부: 작성된 스케줄러 구동 여부

● 차트도구

- 차트도구 메뉴를 클릭하여 창을 연다.
- 차트도구는 시각화 라이브러리인 D3.js 기반으로 시각화 하는 기능을 제공한다.
- 차트를 선택한 후, 입력 경로에서 시각화 할 결과 파일의 경로를 지정한다.
- 목록에서 차트를 선택한다.

★ 차트 목록
click

터링 차트도구 Hive C

워크플로우 스케줄러 진행과정관리 HDFS 브라우저 모니터링 차트도구 Hive Query도구 나의정보 관리자 도구 로그아웃

차트 파라미터

신규 실행 히스토리 다운로드 업로드

클러스터명 noflamingo

차트 목록

파일 경로

첫번째 기록사용 ☐ True ☒ False

구분문자:

차트 제목:

차트 파라미터PDF 그래프

타겟 ID 리스트:

ID 인덱스: +

평균 인덱스: +

기본 인덱스 Dev: +

X-Axis명:

이미지 저장

CL_S_WC
Classification
Intern_test
OptiOutput_201610291814
TV_VOD
TV_VODPreprocess
TV_VOD_UserSimilarity_cf
credit_card_clients_dataset
data
dataman
demo_data
demo_result
expression
ga
home
house
input
iris_keams
knn_test

Preview

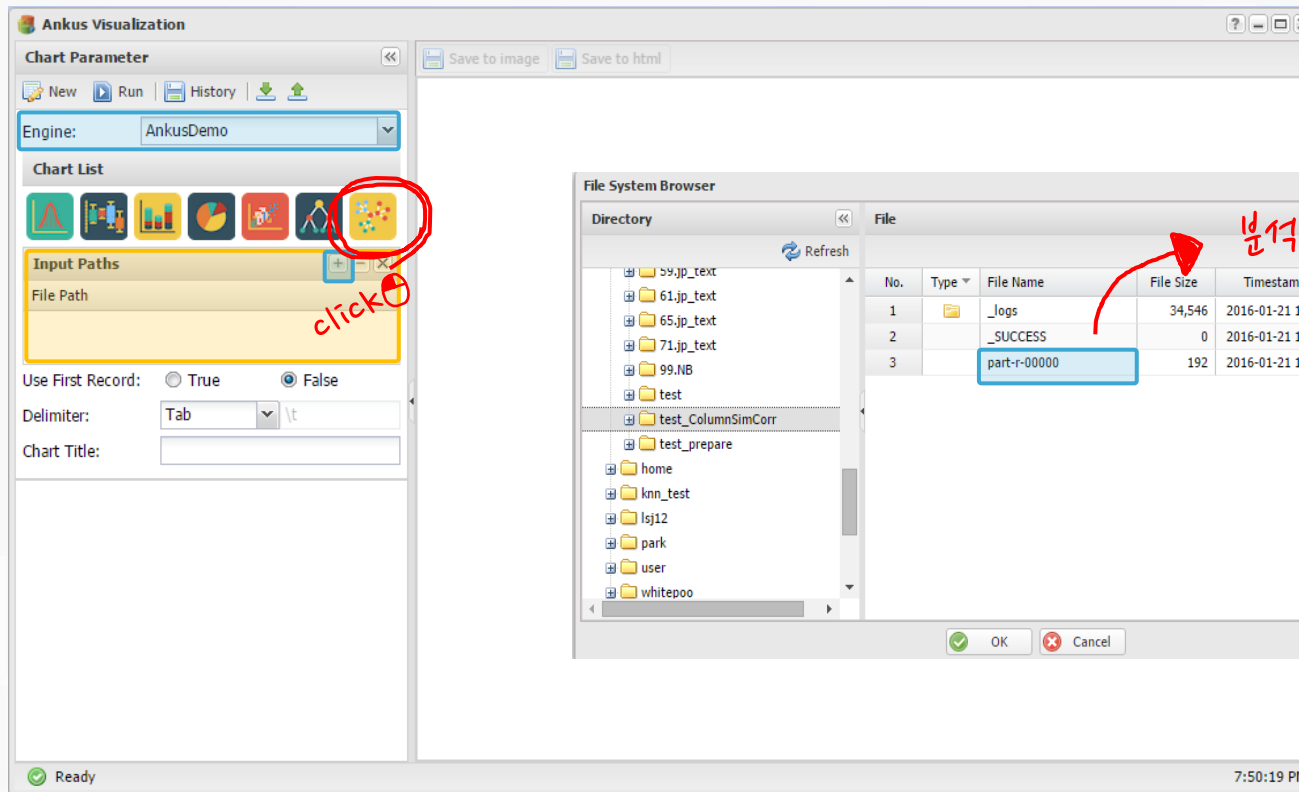
| 유형 | 파일 이름 | 파일 크기 | 최종수정시간 | 파일 접근 권 | 복제 개수 |
|----|------------------------|-----------|------------|-----------|-------|
| | 20160205_115359.jpg | 3316.89kb | 2016-08-11 | rw-r--r-- | 3 |
| | barchart.txt | 0.08kb | 2016-04-26 | rw-r--r-- | 3 |
| | boolean_sample.txt | 0.17kb | 2016-04-29 | rw-r--r-- | 3 |
| | cf_sample.txt | 0.13kb | 2016-04-29 | rw-r--r-- | 3 |
| | cfsum.txt | 0.07kb | 2016-04-29 | rw-r--r-- | 3 |
| | clustering_data_sample | 0.09kb | 2016-04-29 | rw-r--r-- | 3 |
| | contents_sample.txt | 0.12kb | 2016-04-29 | rw-r--r-- | 3 |
| | credit.txt | 148.13kb | 2016-04-27 | rw-r--r-- | 3 |
| | diabetes.txt | 32.97kb | 2016-04-29 | rw-r--r-- | 3 |
| | iris.txt | 4.59kb | 2016-07-18 | rw-r--r-- | 3 |
| | lenses.txt | 0.85kb | 2016-04-29 | rw-r--r-- | 3 |
| | numeric_sample.txt | 0.17kb | 2016-04-29 | rw-r--r-- | 3 |
| | string_sample.txt | 0.03kb | 2016-04-29 | rw-r--r-- | 3 |
| | weather_tennis.txt | 0.37kb | 2016-04-29 | rw-r--r-- | 3 |

분석 결과 파일

● 차트도구

- 차트도구 메뉴를 클릭하여 창을 연다.
- 차트도구는 시각화 라이브러리인 D3.js 기반으로 시각화 하는 기능을 제공한다.
- 차트를 선택한 후, 입력 경로에서 시각화 할 결과 파일의 경로를 지정한다.
- 차트 목록에서 차트를 하나 선택한다.

★ 차트 목록



7:50:19 PM



● 차트 상세 설정

- 선택한 차트의 상세 설정을 입력한다.
- 공통 설정 외에 차트 별로 설정 내용이 상이하므로 데이터를 확인하며 설정한다. ★ 실행하기
- 설정이 끝나면, 실행 버튼으로 시각화를 실행한다.
- 추가적으로, 히스토리는 이전에 실행한 차트를 불러오는 기능이다.

Visualization 히스토리

연진: noflamingo 시작일: 종료일: 상태: 모두 차트 제목: [선택] [초기화]

| 번호 | 차트 제목 | 액션명 | 시작일 | 종료일 | 경과시간 | 진행율 | 상태 | 사용자명 |
|-------|-------------------|-----|---------------------|---------------------|------|------|----|-----------|
| 13846 | hamoni | End | 2017-07-31 09:22:04 | 2017-07-31 09:22:07 | 0.2 | 100% | 성공 | admin |
| 13834 | Visualization Job | End | 2017-07-25 17:04:26 | 2017-07-25 17:04:28 | 0.2 | 100% | 성공 | rockplace |
| 13833 | Visualization Job | End | 2017-07-25 17:03:04 | 2017-07-25 17:03:07 | 0.2 | 100% | 성공 | rockplace |
| 13832 | Visualization Job | End | 2017-07-25 17:01:48 | 2017-07-25 17:01:51 | 0.2 | 100% | 성공 | rockplace |
| 13831 | Visualization Job | End | 2017-07-25 17:01:30 | 2017-07-25 17:01:33 | 0.2 | 100% | 성공 | rockplace |
| 13830 | Visualization Job | End | 2017-07-25 17:01:10 | 2017-07-25 17:01:13 | 0.2 | 100% | 성공 | rockplace |
| 13829 | Visualization Job | End | 2017-07-25 17:00:55 | 2017-07-25 17:01:01 | 0.6 | 100% | 성공 | rockplace |
| 13818 | Visualization Job | End | 2017-07-13 22:29:55 | 2017-07-13 22:29:58 | 0.2 | 100% | 성공 | rockplace |
| 13814 | Visualization Job | End | 2017-06-30 16:27:40 | 2017-06-30 16:27:46 | 0.6 | 100% | 성공 | admin |

[1 / 1] 보기: 1-9 / 9

차트 파라미터

[신규] [실행] [히스토리] [다운로드] [업로드]

클러스터명: noflamingo

차트 목록

입력 경로: [추가] [삭제] [모두 삭제]

파일 경로

• /output/c45_01/C45_rule.txt

첫번째 기록사용 ☐ True ☒ False

구분문자: 탭 |

차트 제목:

분산 플롯

X-Axis 인덱스: +

X-Axis명:

Y-Axis 인덱스: +

Y-Axis명:

클래스 속성 인덱스: +

도트 사이즈 인덱스: +

첫 번째 레코드 반영 여부

입력 데이터 구분문자 선택

chart 이름 입력

X축 인덱스(변수 선택)

X축 이름

Y축 인덱스(변수 선택)

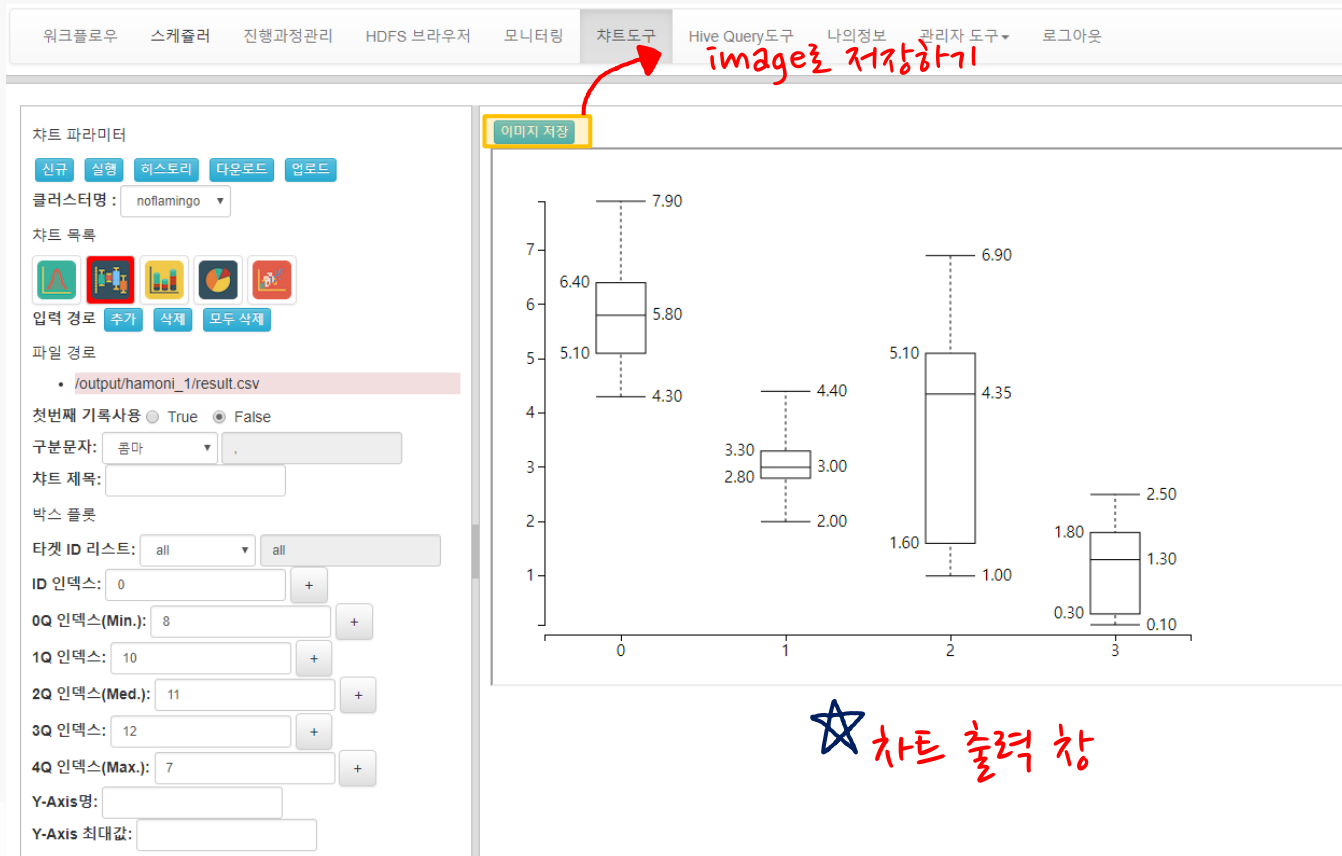
Y축 이름

클래스 속성 인덱스(변수 선택)

도트 사이즈 인덱스(변수 선택)

● 차트 확인

- 실행하면 차트 출력 창에 차트가 나타난다.
- 차트는 image 파일로 저장할 수 있다.

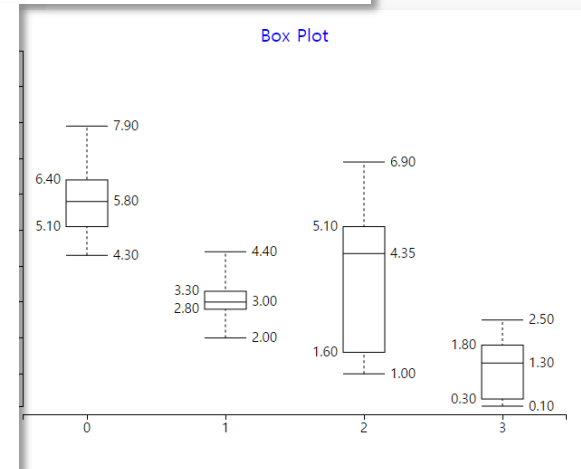
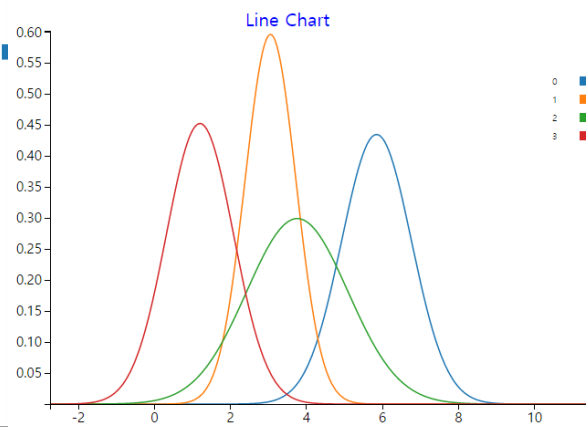
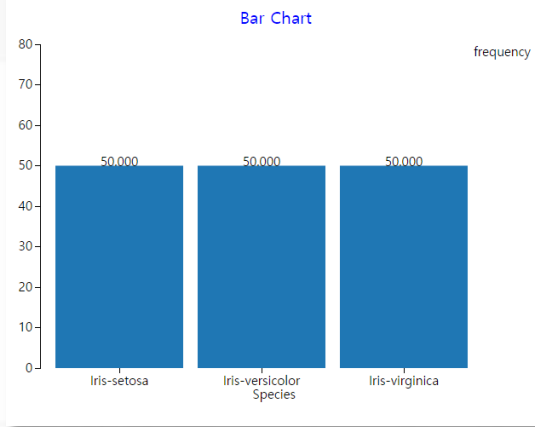
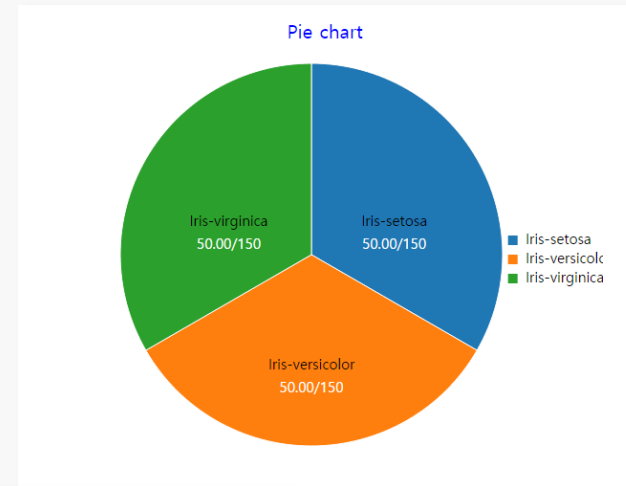
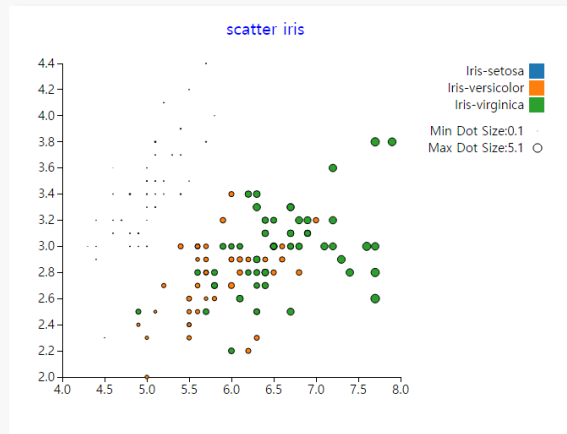


3.1 차트 도구



기타 차트 예시

- ankus Analyzer G 는 현재 버전에서 총 5종의 차트를 제공하고 있다.
(PDF Graph, Box Plot, Bar Chart, Pie Chart, Scatter Plot)
- 분석 결과에 따라 원하는 차트를 그릴 수 있다.



감사합니다.

<http://www.openankus.com>
ankus@onycom.com



ONYCOM