

BM 과정 AI Architecture & Business Model

2019

• 에스튜비즈
대표컨설턴트
이 이 백



목 차

- I. Big Data Architecture와 AI 시스템의 Position
- II. AI 구축 방법론
- III. Industry 4.0
- IV. AI 시스템 구축을 위한 기본 지식과 기술
- V. Deep Learning Revisit
- VI. Business Model과 AI Topic 발굴

IV

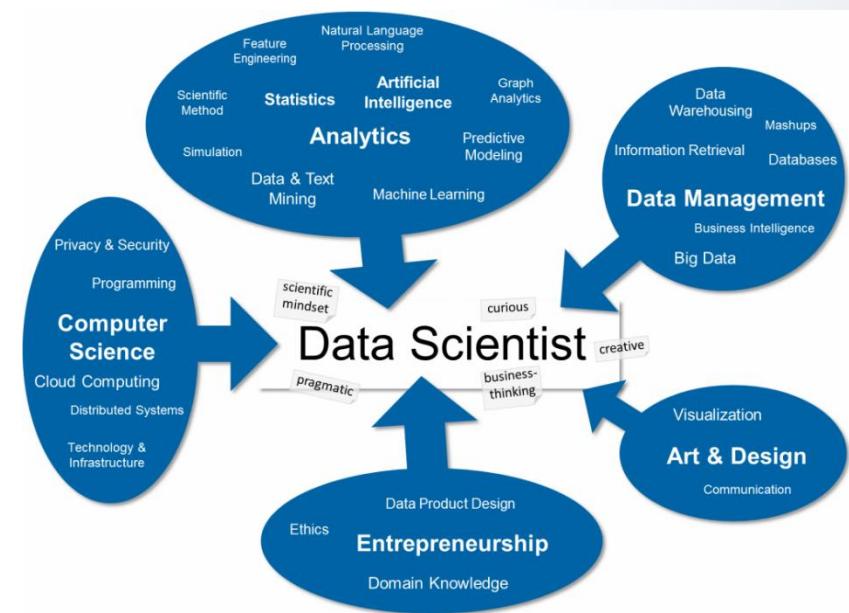
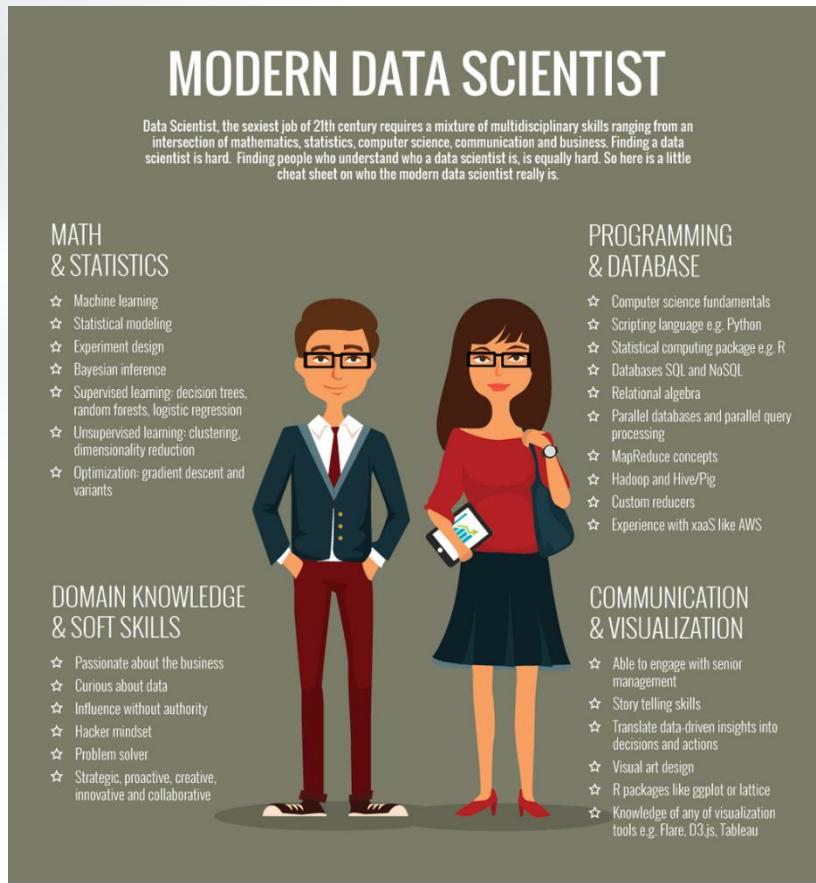
AI 시스템 구축을 위한 기본 지식과 기술

1. Data Scientist의 필요 역량
2. AI Server 설치 및 테스트(실습)
3. 기본 지식과 기술(실습)



1. Data Scientist의 필요 역량

- 빅데이터의 효과적/효율적 활용의 시대적 요구로 비즈니스 도메인 지식과 기술기반 분석을 수행할 수 있는 새로운 역할의 포지션
 - Data Scientist의 부문별 필요 지식 및 기술 아이템**



[source] <https://blog.zhaw.ch/datascience/the-data-science-skill-set/>

1. Data Scientist의 필요 역량

5

■ 자가 역량 평가

* 아래의 자가 평가표는 저자의 도메인 현장 실무, 전략 컨설팅, Stats, ML, DL 분석, Analytic 서비스 구현 경험을 바탕으로 작성한 것임

영 역	내 용	이 해	경 험	구 현
비즈니스 분석	<ul style="list-style-type: none"> ▪ Analytics 서비스 <ul style="list-style-type: none"> • 시장 세분화, 고객등급, 협의탐지, 생산품질분석, 고장진단, 가격예측, Risk Management, 기타 			
비즈니스 실전	<ul style="list-style-type: none"> ▪ 리더 : 해당 업무의 실무 경험형 리더, 조직 경험의 다양성, 실무자 대동 없이 상사와의 실무 심층수주의 정보 전달력, 기타 ▪ 실무자 : 고객점점, 생산 품질, 탐지, 예측, 진단 등의 5년 이상의 심층 체험 ▪ 컨설턴트 : 200페이지 이상의 현장 진단/전략 보고서 작성, C-Level 발표, 기타 			
BI Tool	<ul style="list-style-type: none"> ▪ OLAP Tool, Dashboard, Report, Script Coding, 			
DBMS	<ul style="list-style-type: none"> ▪ Database Modeling, Table Creation, Schema/Meta, Performance Tuning, Stored Procedure, SQL, 기타 			
Big Data	<ul style="list-style-type: none"> ▪ Open Source Community, Hadoop, Eco-system, YARN, Mahout, Hive, Hbase, Kafka, Pig, Zookeeper, Flume, Spark, AWS, Azure, Google Cloud 			
Cloud	<ul style="list-style-type: none"> ▪ AWS, Azure Gcloud, Other Cloud 			
Data Processing	<ul style="list-style-type: none"> ▪ Stored Procedure, ODBC/JDBC 연계 프로그래밍, ▪ SQL : 다중 join, group by, order by, sub-sql, union, SQL Function(decode, nvl, to_char, to_num, timestamps, etc), create, insert, update, delete, etc 			

1. Data Scientist의 필요 역량

6

■ 자가 역량 평가

* 아래의 자가 평가표는 저자의 도메인 현장 실무, 전략 컨설팅, Stats, ML, DL 분석, Analytic 서비스 구현 경험을 바탕으로 작성한 것임

영 역	내 용	이 해	경 험	구 현
Language	<ul style="list-style-type: none"> ▪ Assembler, C, C++, Java, Scala, Python, JavaScript, R, 기타 			
Mathematics	<ul style="list-style-type: none"> ▪ Vector, Matric, Linear Transformation, Linear Algebra, Differential, Integral 			
Statistics	<ul style="list-style-type: none"> ▪ Tukey, Bonferroni, Eigen Vector, Bernoulli, Gaussian, Laplace, Winsorization 			
Machine Learning	<ul style="list-style-type: none"> ▪ kNN, Vapnik, Naïve Bayesian, Hinge Loss, Random Block, Bagging, Sparse, scikit-learn, 			
Deep Learning	<ul style="list-style-type: none"> ▪ Debugging, Feature Engineering, Adams, Dropout, Hypothesis, Batch Normalization, Autoencoder, GAN, Hinton, Q learning, DeepLearning4j, MXNET 			
Web	<ul style="list-style-type: none"> ▪ REST API, React, Angular, Spring, NodeJs, V8, npm, ng serve, wget 			
OS / shell	<ul style="list-style-type: none"> ▪ Kernel, ssh, Protocol 2, MPP, systemctl, scp, NAT, bash, curl 			
Project & Library	<ul style="list-style-type: none"> ▪ Maven, pom.xml, build.sbt, gradle, GitHub, Docker 			

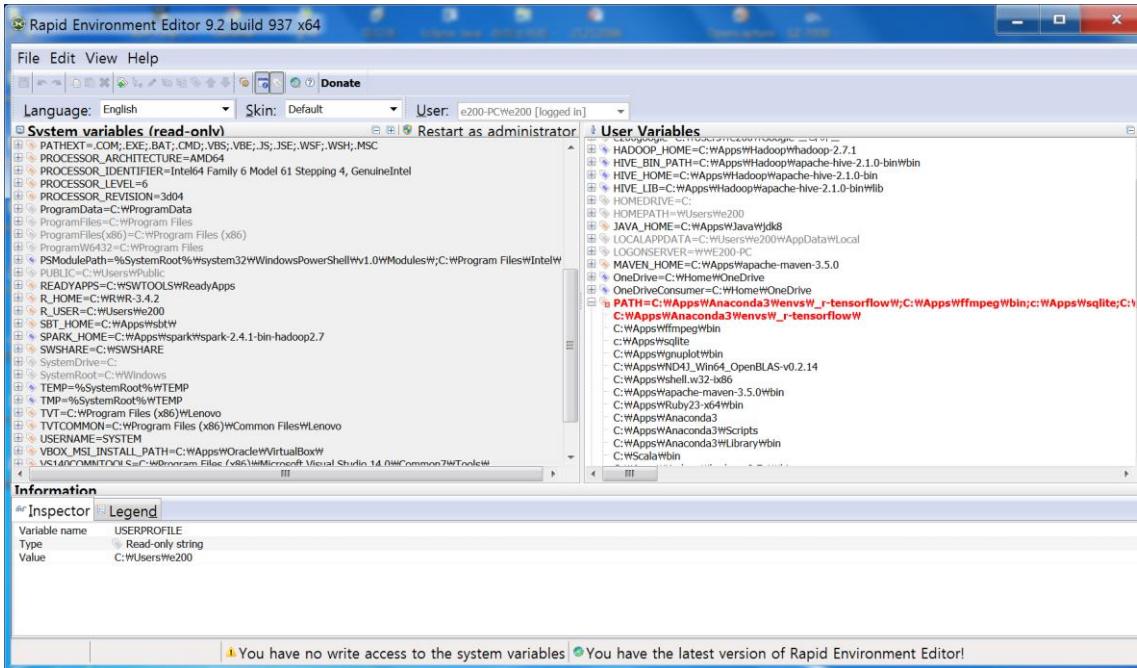
1. Data Scientist의 필요 역량

- 요청자와의 소통 능력 – 요청자 Business Domain에 대한 정확한 이해
- Consultancy
- Language
 - C, C++, Go
 - Java, Scala
 - Python, R
 - Javascript, HTML, CSS
- Data Repository
 - DBMS
 - HADOOP
- Mathematics
 - 백터, 행렬, 텐서
 - 미분
 - Mathematical Function
- Data Processing
- Statistics, Machine Learning, Deep Learning

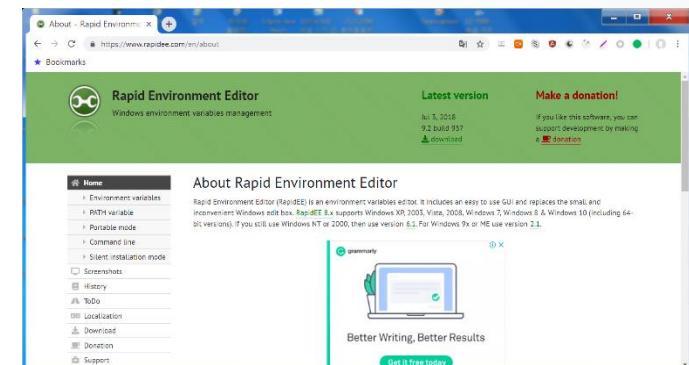
2. AI Server 설치 및 테스트(실습)

▪ Rapid Environment Editor

- 빅데이터 관련 소프트웨어 설치 이전에 Windows PC의 설치 상태 확인
- 사용하고자 하는 소프트웨어가 이미 설치되어 있는지 확인
- 동일 소프트웨어의 다른 버전을 설치할 때 충돌 방지를 위하여 환경변수 참조
- Windows7이 제공하는 도구의 불편성 해소를 위해 유용한 툴(rapidEE) 설치
- JAVA_HOME 환경 변수 유통 Check
 - ✓ JAVA_HOME=C:\any_directory\any_path



<https://www.rapidee.com/en/about>



2. AI Server 설치 및 테스트(실습)

▪ 가상OS설치 도구

- 실습 : Windows에 CentOS 7(Linux 배포판) 설치
- CentOS 서버 서비스를 받기 위한 설정
 - ✓ 외부망 연결 = 어댑터1 NAT 포트 22
 - ✓ 내부망 연결 = 어댑터2 호스트전용어댑터

<https://www.virtualbox.org/>



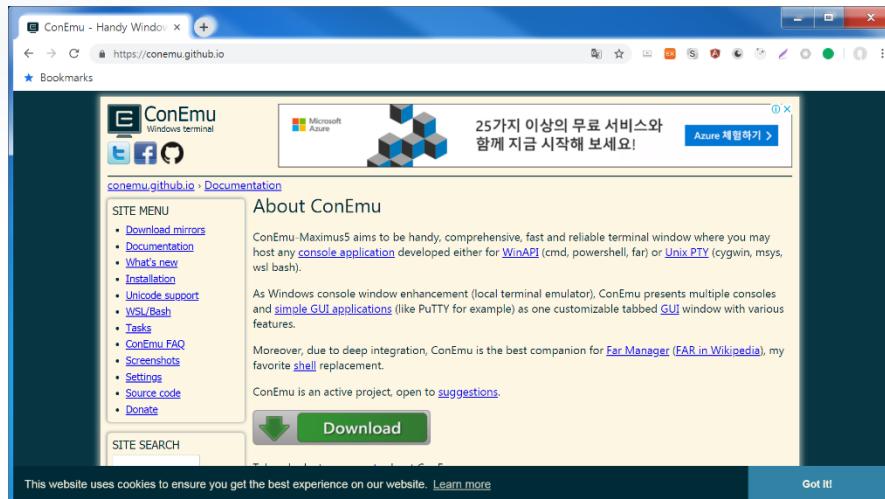
2. AI Server 설치 및 테스트(실습)

10

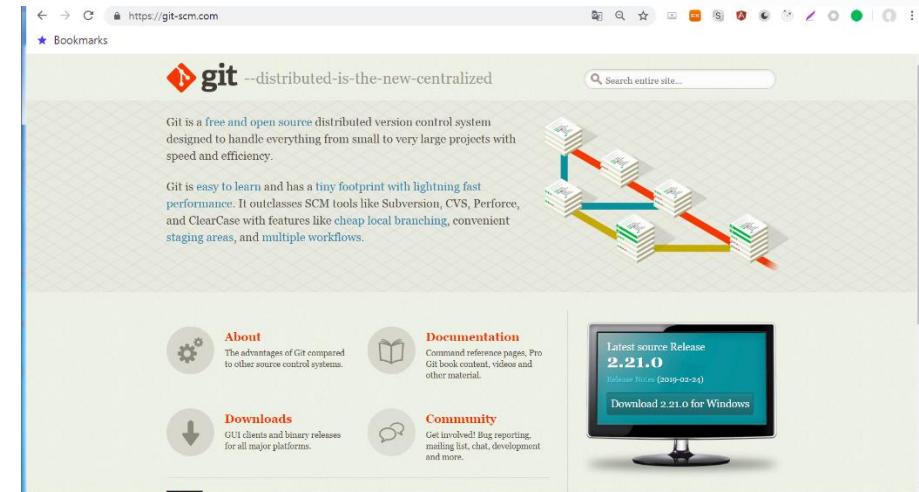
▪ ConEmu

- Windows와 Linux 간 원활한 소통 도구
 - ✓ Ubuntu Server Computer에서 작업은 불편하기도 하며, 위험 유발
- Windows Client에서 콘솔을 이용하여 리눅스 서버에 접속
- Conemu
- Conemu와 git-scm 연동
 - ✓ Windows OS command를 Linux bash 처럼 사용

<https://conemu.github.io/>



<https://git-scm.com/>



2. AI Server 설치 및 테스트(실습)

11

- Linux Server 기본 정보 Check

```
(base)
e200@e200-PC MINGW64 ~
$ ssh e200@169.254.50.164
e200@169.254.50.164's password: □
```

```
Last login: Mon Apr 29 09:21:00 2019 from gateway
[root@localhost ~]# whoami
root
[root@localhost ~]# who am i
e200      pts/0          2019-04-29 16:36 (gateway)
[root@localhost ~]# □
```

```
[root@localhost ~]# ping www.google.com
PING www.google.com (216.58.199.4) 56(84) bytes of data.
64 bytes from hkg12s02-in-f4.1e100.net (216.58.199.4): icmp_seq=1 ttl=47 time=100 ms
64 bytes from hkg12s02-in-f4.1e100.net (216.58.199.4): icmp_seq=2 ttl=47 time=101 ms
64 bytes from hkg12s02-in-f4.1e100.net (216.58.199.4): icmp_seq=3 ttl=47 time=100 ms
^C
--- www.google.com ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2004ms
rtt min/avg/max/mdev = 100.558/101.139/101.930/0.685 ms
```

2. AI Server 설치 및 테스트(실습)

12

- Linux Server 기본 정보 Check

```
[root@localhost ~]# ip addr
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1
000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
        valid_lft forever preferred_lft forever
    inet6 ::1/128 scope host
        valid_lft forever preferred_lft forever
2: enp0s3: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP group default qlen 1000
    link/ether 08:00:27:66:17:bf brd ff:ff:ff:ff:ff:ff
    inet 10.0.2.15/24 brd 10.0.2.255 scope global noprefixroute dynamic enp0s3
        valid_lft 58738sec preferred_lft 58738sec
    inet6 fe80::e10:3dd6:ac1f:cf59/64 scope link noprefixroute
        valid_lft forever preferred_lft forever
3: enp0s8: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP group default qlen 1000
    link/ether 08:00:27:ab:f8:c0 brd ff:ff:ff:ff:ff:ff
    inet 169.254.50.164/24 brd 169.254.50.255 scope link noprefixroute enp0s8
        valid_lft forever preferred_lft forever
    inet 192.168.0.28/24 brd 192.168.0.255 scope global enp0s8
        valid_lft forever preferred_lft forever
    inet6 fe80::a00:27ff:feab:f8c0/64 scope link
        valid_lft forever preferred_lft forever
[root@localhost ~]# ip addr
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
        valid_lft forever preferred_lft forever
    inet6 ::1/128 scope host
        valid_lft forever preferred_lft forever
2: enp0s3: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP group default qlen 1000
    link/ether 08:00:27:66:17:bf brd ff:ff:ff:ff:ff:ff
    inet 10.0.2.15/24 brd 10.0.2.255 scope global noprefixroute dynamic enp0s3
        valid_lft 58697sec preferred_lft 58697sec
    inet6 fe80::e10:3dd6:ac1f:cf59/64 scope link noprefixroute
        valid_lft forever preferred_lft forever
3: enp0s8: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP group default qlen 1000
    link/ether 08:00:27:ab:f8:c0 brd ff:ff:ff:ff:ff:ff
    inet 169.254.50.164/24 brd 169.254.50.255 scope link noprefixroute enp0s8
        valid_lft forever preferred_lft forever
    inet 192.168.0.28/24 brd 192.168.0.255 scope global enp0s8
        valid_lft forever preferred_lft forever
    inet6 fe80::a00:27ff:feab:f8c0/64 scope link
        valid_lft forever preferred_lft forever
[root@localhost ~]#
```

```
[root@localhost ~]# pwd
/home/e200
[root@localhost ~]# []
```

```
/home/e200
[root@localhost ~]# ls
Apps
[root@localhost ~]# []
```

```
[root@localhost ~]# ls -al
total 28
drwx----- 4 1000 e200 123 Apr 13 04:17 .
drwxr-xr-x  3 root root 18 Apr 13 00:25 ..
drwxrwxr-x  5 1000 e200 154 Apr 13 04:14 Apps
-rw-----  1 1000 e200 4936 Apr 29 16:33 .bash_history
-rw-r--r--  1 1000 e200 18 Oct 30 13:07 .bash_logout
-rw-r--r--  1 1000 e200 193 Oct 30 13:07 .bash_profile
-rw-r--r--  1 1000 e200 310 Apr 13 01:18 .bashrc
drwx----- 2 1000 e200 116 Apr 13 02:48 .ssh
-rw-----  1 1000 e200 6095 Apr 13 04:17 .viminfo
[root@localhost ~]# []
```

2. AI Server 설치 및 테스트(실습)

13

▪ Linux Server 기본 정보 Check

```
[root@localhost ~]# yum install openssh-server openssh-client openssh-askpass]
```

```
[root@localhost ~]#
[root@localhost ~]# cat /etc/ssh/sshd_config [
```

```
[root@localhost ~]# cat /etc/ssh/sshd_config
# $OpenBSD: sshd_config,v 1.100 2016/08/15 12:32:04 naddy Exp $
# This is the sshd server system-wide configuration file. See
# sshd_config(5) for more information.

# This sshd was compiled with PATH=/usr/local/bin:/usr/bin

# The strategy used for options in the default sshd_config shipped with
# OpenSSH is to specify options with their default value where
# possible, but leave them commented. Uncommented options override the
# default value.

# If you want to change the port on a SELinux system, you have to tell
# SELinux about this change.
# semanage port -a -t ssh_port_t -p tcp #PORTNUMBER
#
# Port 22
#AddressFamily any
#ListenAddress 0.0.0.0
#ListenAddress :::
# Protocol 2, 1
Protocol 2
HostKey /etc/ssh/ssh_host_rsa_key
#HostKey /etc/ssh/ssh_host_dsa_key
HostKey /etc/ssh/ssh_host_ecdsa_key
HostKey /etc/ssh/ssh_host_ed25519_key

# Ciphers and keying
#RekeyLimit default none

# Logging
#SyslogFacility AUTH
SyslogFacility AUTHPRIV
#LogLevel INFO

# Authentication:

#LoginGraceTime 2m
#PermitRootLogin yes
#StrictModes yes
#MaxAuthTries 6
#MaxSessions 10

#PubkeyAuthentication yes
PubkeyAuthentication yes

# The default is to check both .ssh/authorized_keys and .ssh/authorized_keys2
# but this is overridden so installations will only check .ssh/authorized_keys
AuthorizedKeysFile      .ssh/authorized_keys

#AuthorizedPrincipalsFile none

#AuthorizedKeysCommand none
#AuthorizedKeysCommandUser nobody

# For this to work you will also need host keys in /etc/ssh/ssh_known_hosts
#HostbasedAuthentication no
# Change to yes if you don't trust ~/.ssh/known_hosts for
# HostbasedAuthentication
#IgnoreUserKnownHosts no
# Don't read the user's ~/.rhosts and ~/.shosts files
#IgnoreRhosts yes

# To disable tunneled clear text passwords, change to no here!
#PasswordAuthentication yes
#PermitEmptyPasswords no
PasswordAuthentication yes

# Change to no to disable s/key passwords
#ChallengeResponseAuthentication yes
ChallengeResponseAuthentication no

# Kerberos options
#KerberosAuthentication no
#KerberosOrLocalPasswd yes
#KerberosTicketCleanup yes
#KerberosGetAFSToken no
#KerberosUseKuserok yes

# GSSAPI options
GSSAPIAuthentication yes
GSSAPICleanupCredentials no
#GSSAPIStrictAcceptorCheck yes
#GSSAPIKeyExchange no
#GSSAPIEnable5users no

# Set this to 'yes' to enable PAM authentication, account processing,
# and session processing. If this is enabled, PAM authentication will
# be allowed through the ChallengeResponseAuthentication and
# PasswordAuthentication. Depending on your PAM configuration,
# PAM authentication via ChallengeResponseAuthentication may bypass
# the setting of "PermitRootLogin without-password".
# If you just want the PAM account and session checks to run without
# PAM authentication, then enable this but set PasswordAuthentication
# and ChallengeResponseAuthentication to 'no'.
# WARNING: 'UsePAM no' is not supported in Red Hat Enterprise Linux and may cau
# problems.
UsePAM yes

#AllowAgentForwarding yes
#AllowTcpForwarding yes
#GatewayPorts no
X11Forwarding yes
#X11DisplayOffset 10
#X11UseLocalhost yes
#PermitTTY yes
#PrintMotd yes
#PrintLastLog yes
#TCPKeepAlive yes
#UseLogin no
#UsePrivilegeSeparation sandbox
#PermitUserEnvironment no
#Compression delayed
#ClientAliveInterval 0
#ClientAliveCountMax 3
#ShowPatchLevel no
#UseDNS yes
#PidFile /var/run/sshd.pid
#MaxStartups 10:30:100
#PermitTunnel no
#ChrootDirectory none
#VersionAddendum none

# no default banner path
#Banner none

# Accept locale-related environment variables
AcceptEnv LANG LC_CTYPE LC_NUMERIC LC_TIME LC_COLLATE LC_MONETARY LC_MESSAGES
```

2. AI Server 설치 및 테스트(실습)

14

▪ Linux Server SSH 설정

- Client(예:Windows)와 Server 간의 원활한 소통
- note
 - 방화벽 해제는 교육 실습의 효율성을 위한 임시 조치임.
 - 실제 환경에서는 방화벽을 활성화하고, 사용자 및 Port 설정을 엄격하게 통제해야 함.

```
[root@localhost ~]#  
[root@localhost ~]# vi /etc/ssh/sshd_config [ ]  
  
# semanage port -a -t ssh_port_t -p tcp #PORTNUMBER  
#  
# Port 22
```

```
# semanage port -a -t ssh_port_t -p tcp #PORTNUMBER  
#  
Port 22
```

```
[root@localhost ~]# firewall-cmd --permanent --zone=public --add-port=22/tcp  
FirewallD is not running
```

```
[root@localhost ~]#  
[root@localhost ~]# reboot[]
```

2. AI Server 설치 및 테스트(실습)

- vi / vim 단축기



[source] <https://nolboo.kim/blog/2016/11/15/vim-for-beginner/>

2. AI Server 설치 및 테스트(실습)

▪ Linux Server 기본 명령어

```
[root@localhost ~]# cat .bashrc
# .bashrc
```

```
# Source global definitions
if [ -f /etc/bashrc ]; then
    . /etc/bashrc
fi

# Uncomment the following line if you don't like systemctl's auto-paging feature:
# export SYSTEMD_PAGER=
```

```
# User specific aliases and functions
export JAVA_HOME=/home/e200/Apps/jdk1.8.0_201
export PATH=$PATH:$JAVA_HOME/bin
```

```
[root@localhost ~]# cat .bashrc
# .bashrc
```

```
# Source global definitions
if [ -f /etc/bashrc ]; then
    . /etc/bashrc
fi
```

```
# Uncomment the following line if you don't like systemctl's auto-paging feature:
# export SYSTEMD_PAGER=
```

```
# User specific aliases and functions
export JAVA_HOME=/home/e200/Apps/jdk1.8.0_201
export PATH=$PATH:$JAVA_HOME/bin
```

```
[root@localhost ~]# []
```

```
[root@localhost ~]# scp [-12346BCpqrv] [-c cipher] [-F ssh_config] [-i identity_file]
      [-l limit] [-o ssh_option] [-P port] [-S program]
      [[user@]host1:]file1 ... [[user@]host2:]file2
[root@localhost ~]# []
```

```
[root@localhost ~]# wget
wget: missing URL
Usage: wget [OPTION]... [URL]...

Try `wget --help' for more options.
[root@localhost ~]# []
```

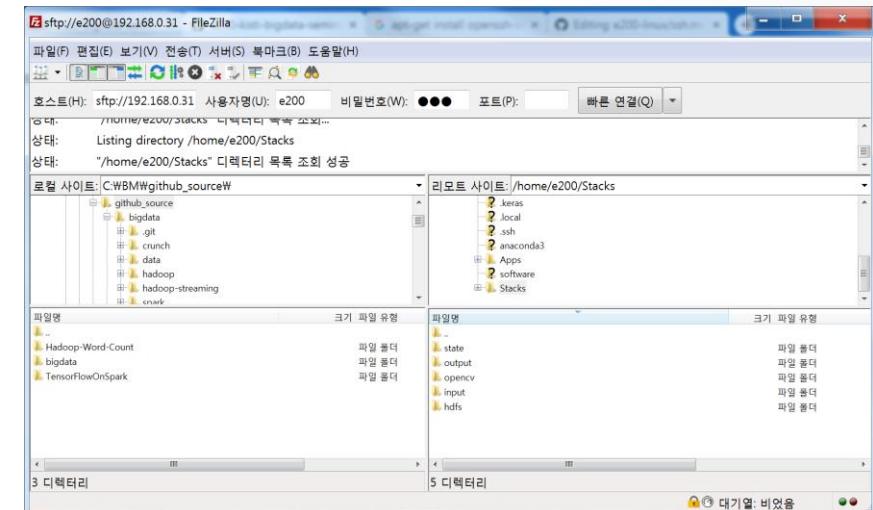
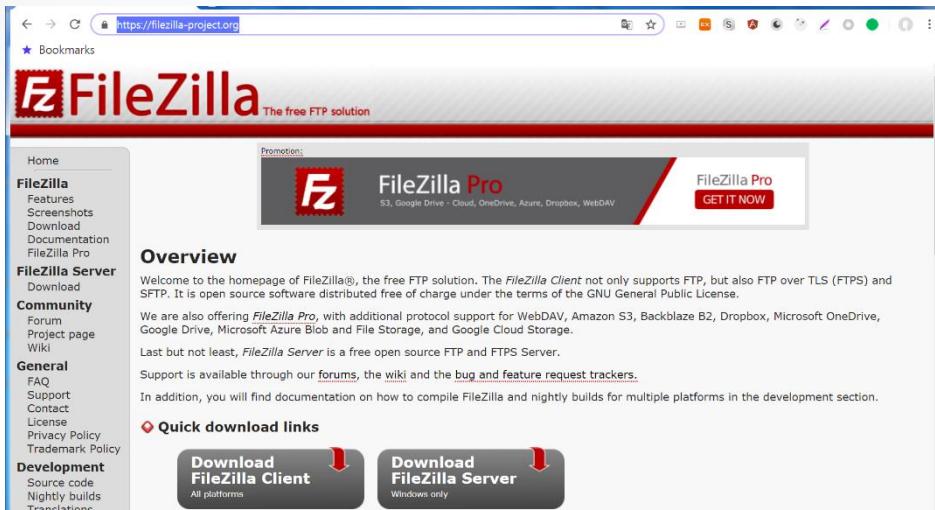
```
[root@localhost ~]# curl
curl: try 'curl --help' or 'curl --manual' for more information
[root@localhost ~]# curl --help
Usage: curl [options...] <url>
Options: (H) means HTTP/HTTPS only, (F) means FTP only
        --anyauth      Pick "any" authentication method (H)
        -a, --append    Append to target file when uploading (F/SFTP)
        --basic        Use HTTP Basic Authentication (H)
        --cacert FILE  CA certificate to verify peer against (SSL)
        --capath DIR   CA directory to verify peer against (SSL)
        -E, --cert CERT[:PASSWD] Client certificate file and password (SSL)
        --cert-type TYPE Certificate file type (DER/PEM/ENG) (SSL)
        --ciphers LIST  SSL ciphers to use (SSL)
        --compressed   Request compressed response (using deflate or gzip)
        -K, --config FILE Specify which config file to read
        --connect-timeout SECONDS Maximum time allowed for connection
        -C, --continue-at OFFSET Resumed transfer offset
        -b, --cookie STRING/FILE String or file to read cookies from (H)
        -c, --cookie-jar FILE Write cookies to this file after operation (H)
        --create-dirs   Create necessary local directory hierarchy
        --crlf        Convert LF to CRLF in upload
        --crlfile FILE Get a CRL list in PEM format from the given file
        -d, --data DATA  HTTP POST data (H)
        --data-ascii DATA HTTP POST ASCII data (H)
        --data-binary DATA HTTP POST binary data (H)
        --data-urlencode DATA HTTP POST data url encoded (H)
        --delegation STRING GSS-API delegation permission
        --digest       Use HTTP Digest Authentication (H)
        --disable-eprt Inhibit using EPRT or LPRT (F)
        --disable-epsv Inhibit using EPSV (F)
        -D, --dump-header FILE Write the headers to this file
        --egd-file FILE EGD socket path for random data (SSL)
        --engine ENGINE  Crypto engine (SSL). "--engine list" for list
        -f, --fail      Fail silently (no output at all) on HTTP errors (H)
        -F, --form CONTENT Specify HTTP multipart POST data (H)
        --form-string STRING Specify HTTP multipart POST data (H)
        --ftp-account DATA Account data string (F)
        --ftp-alternative-to-user COMMAND String to replace "USER [name]" (F)
        --ftp-create-dirs Create the remote dirs if not present (F)
        --ftp-method [MULTICWD/NOCWD/SINGLECWD] Control CWD usage (F)
        --ftp-pasv     Use PASV/EPSV instead of PORT (F)
        -P, --ftp-port ADR Use PORT with given address instead of PASV (F)
        --ftp-skip-pasv-ip Skip the IP address for PASV (F)
        --ftp-pret     Send PRET before PASV (for drftpd) (F)
        --ftp-ssl-ccc  Send CCC after authenticating (F)
```

2. AI Server 설치 및 테스트(실습)

▪ Filezilla

- sftp
- Linux와 Window 간의 파일 교환
- 유통성
 - Windows에서 작성한 코드, 파일을 Linux로 전송
 - 리눅스에서 구현한 코드를 윈도우 딜렉토리에 Backu-Up
 - Virtual Box를 이용한 가상 Linux 개발 중에 vdi 파일에 손상이 가게되면 Linux에서 구현한 코드의 복구 불가할 수 있으니, 주기적 Backup 필수

<https://filezilla-project.org/>



2. AI Server 설치 및 테스트(실습)

18

▪ Hadoop 설치

- Windows에서 Hadoop 설치는 불가능한 것은 아니나, 매우 험난한 길

The screenshot shows a web browser window with the title bar "Apache Hadoop". The address bar contains the URL "https://hadoop.apache.org". The page content is the Apache Hadoop homepage. It features the Apache Hadoop logo (a yellow elephant) and the text "APACHE hadoop". Below the logo, a sub-headline reads "Apache Hadoop". The main text on the page describes the project as developing open-source software for reliable, scalable, distributed computing. It explains that the software library allows for the distributed processing of large data sets across clusters of computers using simple programming models. The text highlights that the library is designed to scale up from single servers to thousands of machines, each offering local computation and storage. It also mentions that the library is designed to detect and handle failures at the application layer, ensuring a highly-available service. At the bottom of the page, there are three buttons: "Learn more »" (blue), "Download »" (orange), and "Getting started »" (green). The footer contains links for "Latest news", "Modules", "Related projects", and information about releases and other projects.

Apache Hadoop

https://hadoop.apache.org

Bookmarks

Apache Software Foundation

Apache Hadoop Documentation Community Development Help Old site

APACHE hadoop

Apache Hadoop

The Apache™ Hadoop® project develops open-source software for reliable, scalable, distributed computing.

The Apache Hadoop software library is a framework that allows for the distributed processing of large data sets across clusters of computers using simple programming models. It is designed to scale up from single servers to thousands of machines, each offering local computation and storage. Rather than rely on hardware to deliver high-availability, the library itself is designed to detect and handle failures at the application layer, so delivering a highly-available service on top of a cluster of computers, each of which may be prone to failures.

Learn more » Download » Getting started »

Latest news

Release 3.1.2 available

2019 Feb 6

Modules

The project includes these modules:

Related projects

Other Hadoop-related projects at Apache include:

2. AI Server 설치 및 테스트(실습)

19

▪ Hadoop 설치

- Hadoop Version 선택

The screenshot shows a web browser window with multiple tabs open. The active tab displays the Apache Hadoop download page at <https://hadoop.apache.org/releases.html>. The page header includes the Apache Software Foundation logo and links for Download, Documentation, Community, Development, Help, and Old site. Below the header, a section titled "Download" provides information about the availability of source code and binary tarballs for various Hadoop versions. A table lists the version, release date, source download link, binary download link, and release notes for each version. At the bottom, instructions for verifying releases using GPG are provided.

Version	Release date	Source download	Binary download	Release notes
3.1.2	2019 Feb 6	source (checksum signature)	binary (checksum signature)	Announcement
3.2.0	2019 Jan 16	source (checksum signature)	binary (checksum signature)	Announcement
2.9.2	2018 Nov 19	source (checksum signature)	binary (checksum signature)	Announcement
2.8.5	2018 Sep 15	source (checksum signature)	binary (checksum signature)	Announcement
2.7.7	2018 May 31	source (checksum signature)	binary (checksum signature)	Announcement

To verify Hadoop releases using GPG:

1. Download the release `hadoop-X.Y.Z-src.tar.gz` from a [mirror site](#).
2. Download the signature file `hadoop-X.Y.Z-src.tar.gz.asc` from [Apache](#).
3. Download the [Hadoop KEYS](#) file.
4. `gpg --import KEYS`
5. `gpg --verify hadoop-X.Y.Z-src.tar.gz.asc`

2. AI Server 설치 및 테스트(실습)

20

▪ Hadoop 설치

- wget <http://mirror.navercorp.com/apache/hadoop/common/hadoop-3.2.0/hadoop-3.2.0.tar.gz>

The screenshot shows a web browser window with the Apache Software Foundation's download page for Hadoop 3.2.0. The URL in the address bar is <https://www.apache.org/dyn/closer.cgi/hadoop/common/hadoop-3.2.0/hadoop-3.2.0.tar.gz>. The page features the Apache 20th Anniversary logo on the left and a "CELEBRATING 20 YEARS OF COMMUNITY-LED DEVELOPMENT 'THE APACHE WAY'" banner in the center. On the right, there is a "SUPPORT APACHE" logo. The main content area provides download links for various mirror sites, including:

- HTTP**
 - <http://apache.mirror.cdnetworks.com/hadoop/common/hadoop-3.2.0/hadoop-3.2.0.tar.gz>
 - <http://apache.tt.co.kr/hadoop/common/hadoop-3.2.0/hadoop-3.2.0.tar.gz>
 - <http://mirror.apache-kr.org/hadoop/common/hadoop-3.2.0/hadoop-3.2.0.tar.gz>
 - <http://mirror.navercorp.com/apache/hadoop/common/hadoop-3.2.0/hadoop-3.2.0.tar.gz>

2. AI Server 설치 및 테스트(실습)

- Hadoop 설치
 - Hadoop-3.2.0.tar.gz 파일 확득 여부 확인

```
MINGW64:/c/Users/e200
(base) e200@aiserver:~$ cd Apps/
(base) e200@aiserver:~/Apps$ cd ~;ls -al
total 584
drwxr-xr-x 15 e200 e200 4096 Apr 11 20:43 .
drwxr-xr-x 3 root root 4096 Mar 23 11:15 ..
drwxrwxr-x 26 e200 e200 4096 Mar 24 00:50 anaconda3
drwxrwxr-x 7 e200 e200 4096 Apr 7 10:37 Apps
-rw----- 1 e200 e200 24052 Apr 11 20:08 .bash_history
-rw-r--r-- 1 e200 e200 220 Apr 4 2018 .bash_logout
-rw-r--r-- 1 e200 e200 4702 Apr 7 10:42 .bashrc
drwx----- 3 e200 e200 4096 Apr 6 13:43 .cache
drwxrwxr-x 3 e200 e200 4096 Mar 23 13:10 .conda
drwxrwxr-x 3 e200 e200 4096 Mar 23 13:11 .config
drwx----- 3 e200 e200 4096 Mar 23 11:16 .gnupg
-rw-rw-r-- 1 e200 e200 29910 Apr 11 20:43 hadoop-3.2.0.tar.gz
drwxr-xr-x 5 e200 e200 4096 Apr 6 10:46 .ipython
drwx----- 2 e200 e200 4096 Apr 6 11:47 .jupyter
drwxrwxr-x 2 e200 e200 4096 Mar 23 13:11 .keras
drwx----- 3 e200 e200 4096 Apr 6 10:46 .local
-rw-r--r-- 1 e200 e200 807 Apr 4 2018 .profile
-rw----- 1 e200 e200 5634 Apr 7 05:30 .scala_history
drwxrwxr-x 2 e200 e200 4096 Mar 23 11:58 software
drwx----- 2 e200 e200 4096 Apr 6 14:23 .ssh
drwxrwxr-x 7 e200 e200 4096 Apr 7 03:58 stacks
-rw-r--r-- 1 e200 e200 0 Mar 23 11:25 .sudo_as_admin_successful
-rwrxr-xr-x 1 e200 e200 108 Apr 6 11:44 TFoS_1_spark-standalone-cluster.sh
sshd.exe* [64]:16192
```

« 170402[64] 2/2 [+/-] NUM PRI ‡ 118x26 (25,598) 25V 14648 100%

2. AI Server 설치 및 테스트(실습)

22

- Hadoop 설치
 - Map Reduce 기본 개념 이해

MapReduce Tutorial – Fundamentals of MapReduce with MapReduce Example

101.1K Views



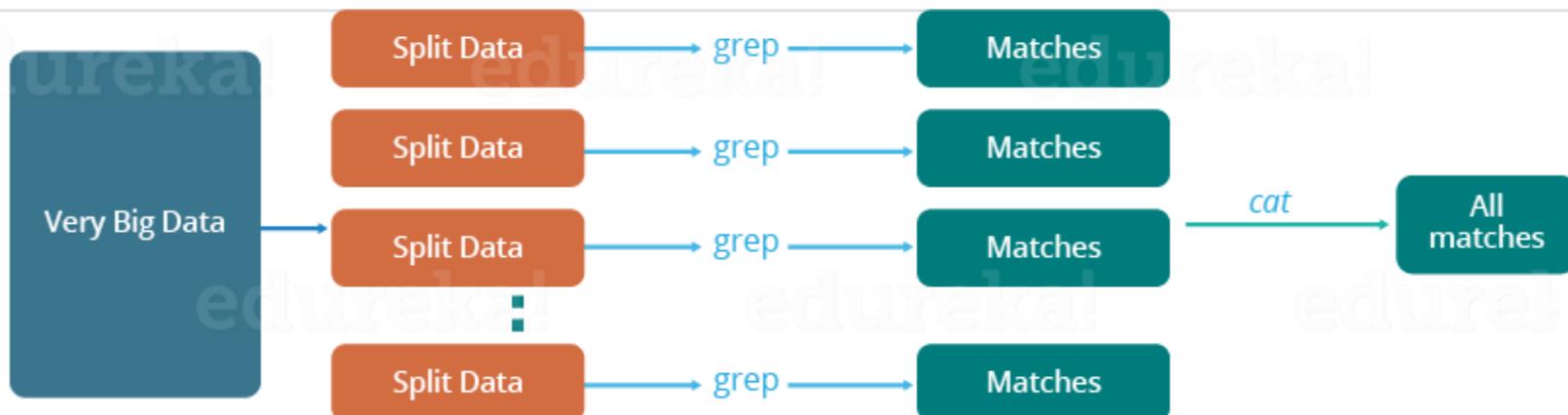
Ashish Bakshi

Published on Dec 25, 2018

The Traditional Way

17 Comments

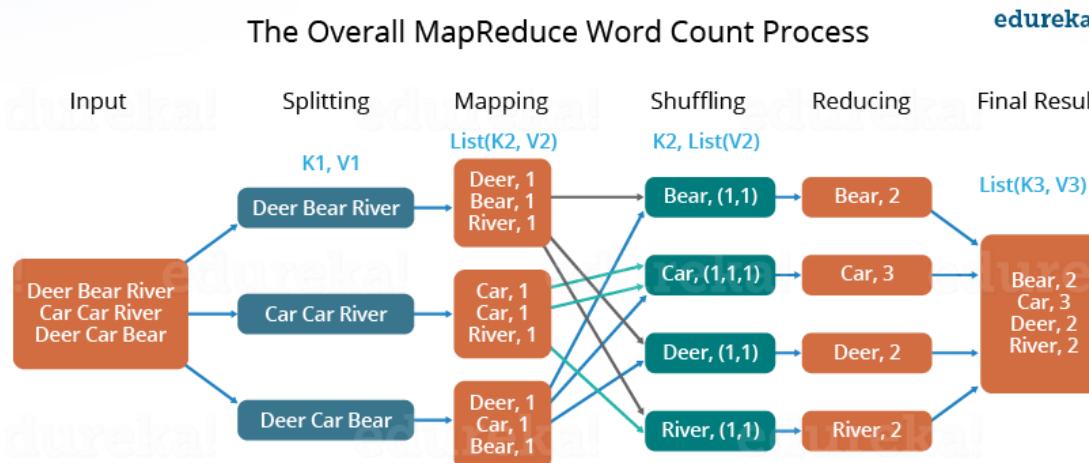
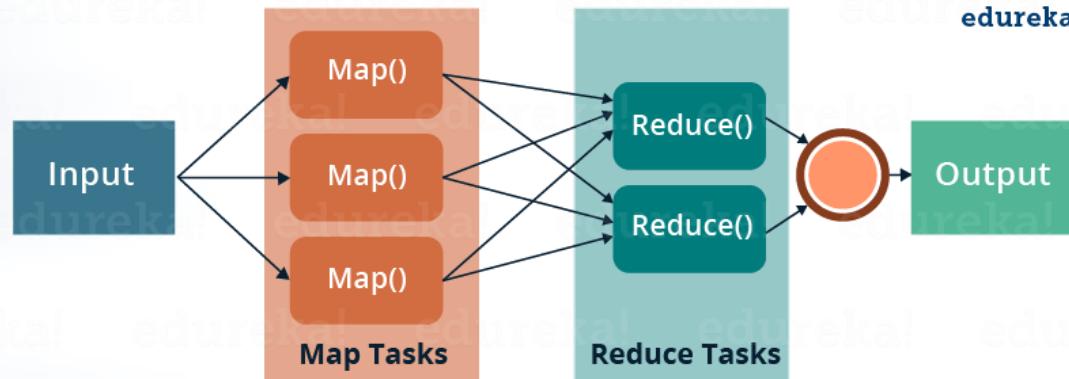
[Bookmark](#)



2. AI Server 설치 및 테스트(실습)

23

- Hadoop 설치
 - Map Reduce 기본 개념 이해



2. AI Server 설치 및 테스트(실습)

24

▪ Map Reduce 체험 실습

- Guide Page

- <https://hadoop.apache.org/docs/r3.2.0/hadoop-project-dist/hadoop-common/SingleCluster.html>

The screenshot shows a web browser window with the following details:

- Title Bar:** Apache Hadoop 3.2.0 - H
- Address Bar:** https://hadoop.apache.org/docs/r3.2.0/hadoop-project-dist/hadoop-common/SingleCluster.html
- Page Content:**
 - Apache Hadoop Logo:** A yellow elephant icon next to the text "APACHE hadoop".
 - The Apache Software Foundation Logo:** A red feather icon with the text "The Apache Software Foundation" and the URL "http://www.apache.org/".
 - Breadcrumbs:** Apache > Hadoop > Apache Hadoop Project Dist POM > Apache Hadoop 3.2.0 > Hadoop: Setting up a Single Node Cluster.
 - Page Statistics:** Wiki | git | Apache Hadoop | Last Published: 2019-01-08 | Version: 3.2.0
 - Section Header:** **Hadoop: Setting up a Single Node Cluster.**
 - Table of Contents:**
 - General
 - Overview
 - Single Node Setup
 - Cluster Setup
 - Commands Reference
 - FileSystem Shell
 - Compatibility
 - Specification
 - Downstream Developer's Guide
 - Admin Compatibility Guide
 - Interface Classification
 - FileSystem Specification
 - Common
 - CLI Mini Cluster
 - Native Libraries
 - Proxy User
 - Rack Awareness
 - Secure Mode
 - Service Level Authorization
 - HTTP Authentication
 - Credential Provider API
 - Hadoop KMS
 - Tracing
 - Unix Shell Guide
 - HDFS
 - Architecture
 - User Guide
 - Purpose:** This document describes how to set up and configure a single-node Hadoop installation so that you can quickly perform simple operations using Hadoop MapReduce and the Hadoop Distributed File System (HDFS).
 - Prerequisites:** Supported Platforms, Required Software, Installing Software.
 - Download:** Available for download.
 - Prepare to Start the Hadoop Cluster:** Standalone Operation, Pseudo-Distributed Operation (Configuration, Setup passphraseless ssh, Execution, YARN on a Single Node), Fully-Distributed Operation.

2. AI Server 설치 및 테스트(실습)

25

▪ Map Reduce 체험 실습

• Map Reduce 작동(1)

```
(base) e200@aiserver:~/Apps/hadoop-3.2.0$  
(base) e200@aiserver:~/Apps/hadoop-3.2.0$ curl https://en.wikipedia.org/wiki/Deep_learning  
(base) e200@aiserver:~/Apps/hadoop-3.2.0$ curl https://en.wikipedia.org/wiki/Deep_learning > ./input/Deep_learning.html  
% Total    % Received % Xferd  Average Speed   Time     Time      Time  Current  
          Dload  Upload   Total Spent  Left Speed  
100  406k  100  406k    0     0  430k      0 --:--:--:--:--:--:--:--  429k
```

```
(base) e200@aiserver:~/Apps/hadoop-3.2.0$ cat ./input/Deep_learning.html | more  
<!DOCTYPE html>  
<html class="client-nojs" lang="en" dir="ltr">  
<head>  
<meta charset="UTF-8"/>  
<title>Deep learning - Wikipedia</title>  
<script>document.documentElement.className = document.documentElement.className.replace( /(^|\s)client-nojs(\s|$)/, "$1client-js$2");</script>  
<script>(window.RLQ=window.RLQ||[]).push(function(){mw.config.set({"wgCanonicalNamespace": "", "wgCanonicalSpecialPageName": false, "wgNamespaceNumber": 0, "wgPageName": "Deep_learning", "wgTitle": "Deep learning", "wgCurRevisionId": 891855721, "wgRevisionId": 891855721, "wgArticleId": 32472154, "wgIsArticle": true, "wgIsRedirect": false, "wgAction": "view", "wgUserName": null, "wgUserGroups": ["*"], "wgCategories": ["CS1 maint: Archived copy as title", "CS1: long volume value", "CS1 maint: BOT: original-url status unknown", "All articles with unsourced statements", "Articles with unsourced statements from April 2018", "Wikipedia articles needing clarification from September 2017", "Articles with unsourced statements from September 2017", "Wikipedia articles that are too technical from July 2016", "All articles that are too technical", "Articles needing expert attention from July 2016", "All articles needing expert attention", "Articles with unsourced statements from July 2016", "Articles prone to spam from June 2015", "Deep learning", "Artificial neural networks", "Artificial intelligence", "Emerging technologies"], "wgBreakFrames": false, "wgPageContentLanguage": "en", "wgPageContentModel": "wikitext", "wgSeparatorTransformTable": [ "", "" ], "wgDigitTransformTable": [ "", "" ], "wgDefaultDateFormat": "dmy", "wgMonthNames": [ "", "January", "February", "March", "April", "May", "June", "July", "August", "September", "October", "November", "December" ], "wgMonthNamesShort": [ "", "Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug", "Sep", "Oct", "Nov", "Dec" ], "wgRelevantPageName": "Deep_learning", "wgRelevantArticleId": 32472154, "wgRequestId": "XK4PqgpAIDEAAFqR-9sAAAAO", "wgCSPNonce": false, "wgIsProbablyEditable": true, "wgRelevantPageIsProbablyEditable": true, "wgRestrictionEdit": [], "wgRestrictionMove": [], "wgFlaggedRevsParams": {"tags": {}}, "wgStableRevisionId": null, "wgCategoryTreePageCategoryOptions": "{\"mode\":0, \"hidetitle\":20, \"showcount\":true, \"namespaces\":false}", "wgBetaFeaturesFeatures": [], "wgMediaViewerOnClick": true, "wgMediaViewerEnabledByDefault": true, "wgPo
```

2. AI Server 설치 및 테스트(실습)

26

▪ Map Reduce 체험 실습

- Map Reduce 작동(2)

```
1121 ls ./input/
1122 cat output/*
1123 curl https://en.wikipedia.org/wiki/Deep_learning
1124 curl https://en.wikipedia.org/wiki/Deep_learning > ./input/Deep_learning.html
1125 cat ./input/Deep_learning.html
1126 cat ./input/Deep_learning.html | more
1127 history
(base) e200@aiserver:~/Apps/hadoop-3.2.0$
```

```
(base) e200@aiserver:~/Apps/hadoop-3.2.0$ bin/hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-3.2.0.jar grep input/Deep_learning.html output/deep 'dee[a-z]+'
2019-04-11 22:27:51,207 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2019-04-11 22:27:51,638 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2019-04-11 22:27:51,639 INFO impl.MetricsSystemImpl: JobTracker metrics system started
2019-04-11 22:27:53,825 INFO input.FileInputFormat: Total input files to process : 1
2019-04-11 22:27:54,003 INFO mapreduce.JobSubmitter: number of splits:1
2019-04-11 22:27:55,067 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local1568987053_0001
2019-04-11 22:27:55,076 INFO mapreduce.JobSubmitter: Executing with tokens: []
2019-04-11 22:27:55,872 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
2019-04-11 22:27:55,874 INFO mapreduce.Job: Running job: job_local1568987053_0001
2019-04-11 22:27:55,892 INFO mapred.LocalJobRunner: OutputCommitter set in config null
2019-04-11 22:27:55,959 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2019-04-11 22:27:55,960 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
2019-04-11 22:27:55,965 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter
2019-04-11 22:27:56,211 INFO mapred.LocalJobRunner: Waiting for map tasks
2019-04-11 22:27:56,219 INFO mapred.LocalJobRunner: Starting task: attempt_local1568987053_0001_m_000000_0
2019-04-11 22:27:56,394 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2019-04-11 22:27:56,396 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
2019-04-11 22:27:56,521 INFO mapred.Task: Using ResourceCalculatorProcessTree : []
2019-04-11 22:27:56,560 INFO mapred.MapTask: Processing split: file:/home/e200/Apps/hadoop-3.2.0/input/Deep_learning.html:0+415899
2019-04-11 22:27:56,920 INFO mapreduce.Job: Job job_local1568987053_0001 running in uber mode : false
```

2. AI Server 설치 및 테스트(실습)

27

- Map Reduce 체험 실습
 - Map Reduce 작동(3)

```
(base) e200@aiserver:~/Apps/hadoop-3.2.0$ cat ./output/deep/*  
169    deep  
5      deeplearning  
4      deepai  
2      deeplearningbook  
2      deeper
```

- Hadoop Map Reduce 실습 두번째
 - Internet에서 관심 Contents 정보 획득
 - Map Reduce 작동
 - 다양한 Word Pattern으로 집계 정보 생성
 - 통계량 집계
 - 실습 완료에 대한 결과 확인 메일 발송
 - Yibeck.Lee@gmail.com

2. AI Server 설치 및 테스트(실습)

28

▪ Spark 설치

- Why? 하둡보다 110배 빠르다;
- 머신러닝과 딥러닝의 구현에 최적
- Deeplearning4j, Tensorflow, MXNET 등 패키지들이 스파크 연동 패키지 제공
- 스칼라로 개발됨
- API 제공 언어 : scala, java, python, R

The screenshot shows the official Apache Spark website at <https://spark.apache.org>. The page features the Apache logo and the text "Lightning-fast unified analytics engine". A navigation bar includes links for Download, Libraries, Documentation, Examples, Community, Developers, and Apache Software Foundation. Below the navigation, a section titled "Speed" compares Hadoop and Spark's performance on a logistic regression task. A bar chart shows Spark's execution time is 0.9 seconds compared to Hadoop's 110 seconds. Another section, "Ease of Use", highlights the ability to write applications in Java, Scala, Python, R, and SQL. A code snippet demonstrates reading JSON data into a DataFrame. The right sidebar contains news items about recent Spark releases and an event for APACHECON North America.

Apache Spark™ is a unified analytics engine for large-scale data processing.

Speed

Run workloads 100x faster.

Apache Spark achieves high performance for both batch and streaming data, using a state-of-the-art DAG scheduler, a query optimizer, and a physical execution engine.

Running time (s)

System	Running time (s)
Hadoop	110
Spark	0.9

Logistic regression in Hadoop and Spark

Ease of Use

Write applications quickly in Java, Scala, Python, R, and SQL.

```
df = spark.read.json("logs.json")
df.where("age > 21")
.select("name,first").show()
```

Spark's Python DataFrame API

Read JSON files with automatic schema inference

Spark offers over 80 high-level operators that make it easy to build parallel apps. You can use it interactively from the Scala, Python, R, and SQL shells.

Latest News

- Spark 2.4.1 released (Mar 31, 2019)
- Spark 2.3.3 released (Feb 15, 2019)
- Spark 2.2.3 released (Jan 11, 2019)
- Spark+AI Summit (April 23-25th, 2019, San Francisco) agenda posted (Dec 19, 2018)

Archive

APACHECON North America Sept. 9-12 Las Vegas, Nevada 2019

Download Spark

Built-in Libraries:

- SQL and DataFrames
- Spark Streaming
- MLlib (machine learning)
- GraphX (graph)
- Third-Party Projects

2. AI Server 설치 및 테스트(실습)

29

- 아나콘다 배포판 설치
 - windows에서 python공식버전(python.org) 설치는 험난한 길
 - 머신러닝 라이브러리인 scikit-learn 사용하려면 visual C 2007 을 구해서 source를 컴파일 해야함
 - 전문가도 짤짤매는게 현실

```
$ mkdir -p ~/Apps/anaconda3
$ curl -O https://repo.anaconda.com/archive/Anaconda3-2019-03-Linux-86.sh
$ sha256sum Anaconda3-2019-03-Linux-86.sh
$ bash Anaconda3-2019-03-Linux-86.sh $ source ~/.bashrc
```

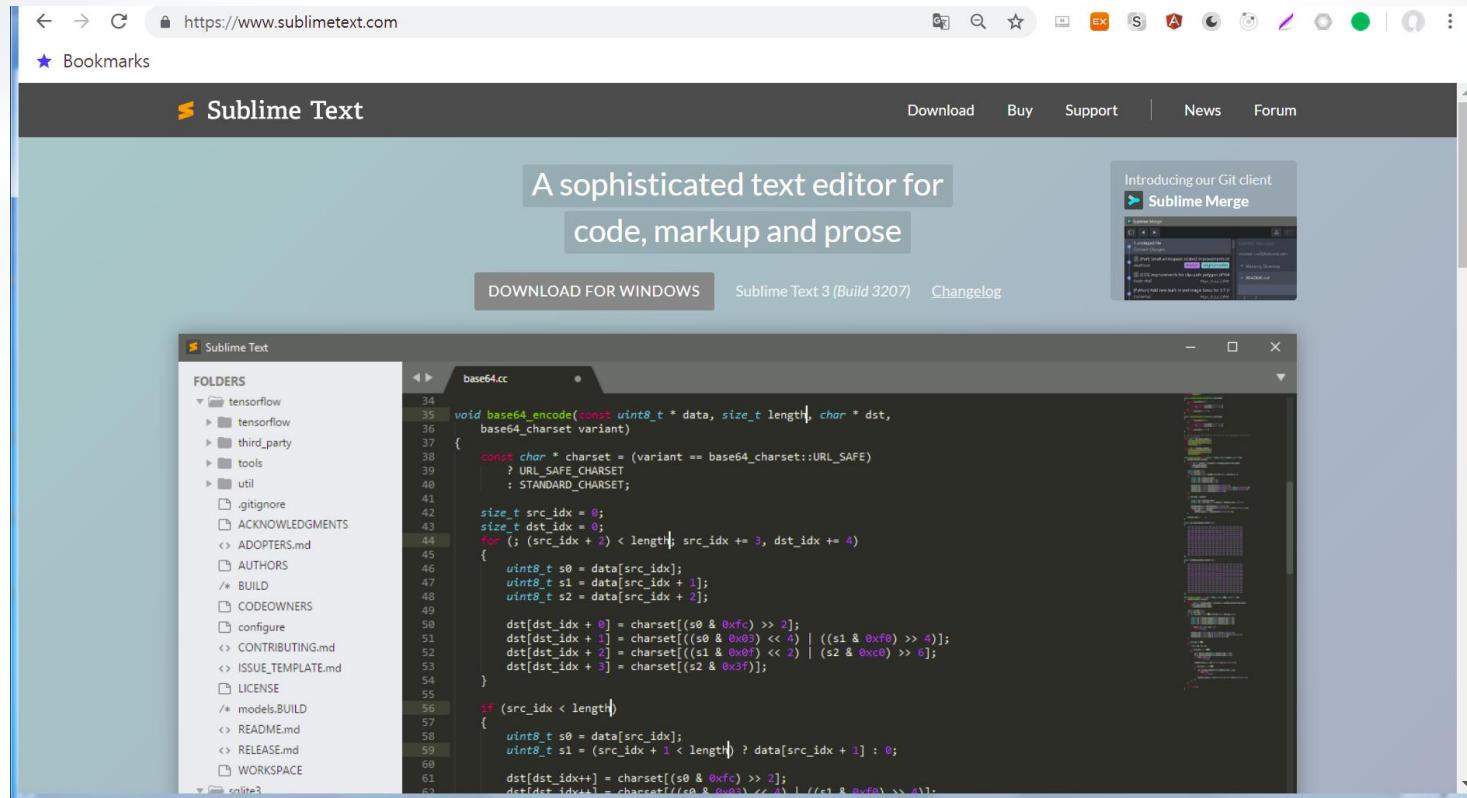
- curl : download in linux
- bash Anaconda3-2019-03-Linux-86.sh : anaconda 설치 shell 실행
- source ~/.bashrc : anaconda을 추가로 인한 사용자 환경 refresh

2. AI Server 설치 및 테스트(실습)

30

▪ 개발 환경의 선택

- 1) IDE : PyCharm, IntelliJ, Eclipse,
- 2) Programming 전용 Editor : sublime, atom, editplus, notepad++, ...
- 편집기는 Sublime으로 기본으로 사용하나, 교육생은 자신에게 익숙한 편집기를 사용하여도 무방함

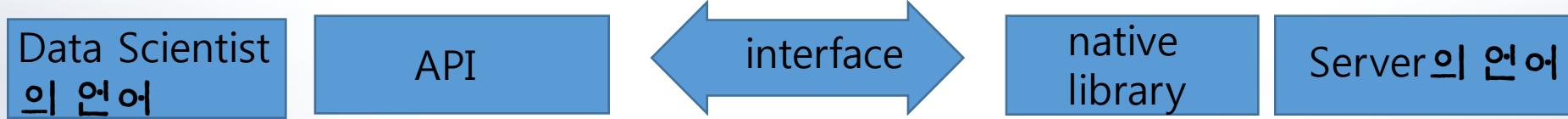


3. 기본 지식과 기술(실습)

31

- Competence 1. Language와 API 선택

- 적절한 컴퓨터 언어를 선택하고, 제공하는 API(Application Program Interface) Call 하여 분석 작업을 수행함



C/C++ include math.h

java

import math

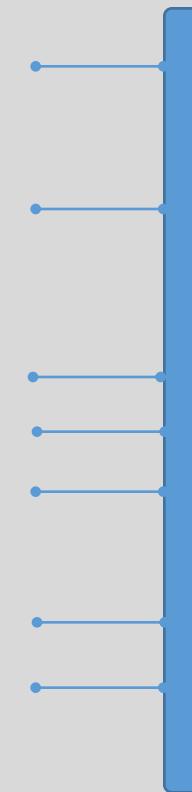
scala

Python

import numpy
import sklearn
import tensorflow

R

library(rpart)
library(mxnet)



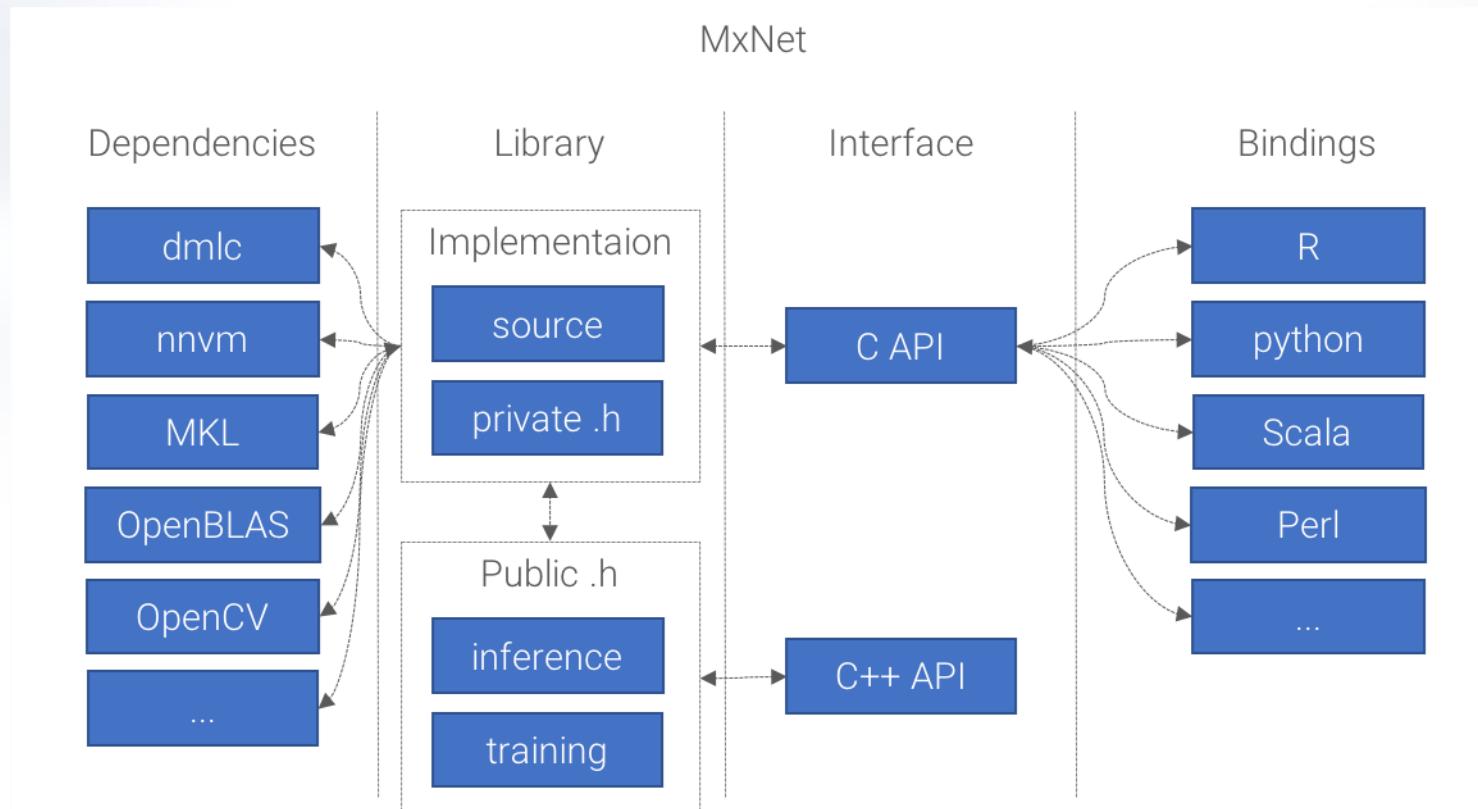
C
C++

Java

3. 기본 지식과 기술(실습)

32

- Language



[source] <https://www.jamesserra.com/archive/2015/08/relational-databases-vs-non-relational-databases/>

3. 기본 지식과 기술(실습)

33

- Database

		Relational	Non-Relational
Analytics	Proprietary Storage	Amazon Redshift EMC Greenplum HP Vertica	IBM Netezza Oracle Teradata MPP
	Hadoop Storage	Cloudera Impala Presto	Hive SQL-on-Hadoop
Operational	Proprietary Storage	Traditional SQL	NewSQL
		Oracle DB2 SQL Server MySQL	User-Sharded MySQL NuoDB Clustrix On-Disk MemSQL VoltDB In-Memory
	Hadoop Storage		Splice Machine On-Hadoop
			Column Family: HBase

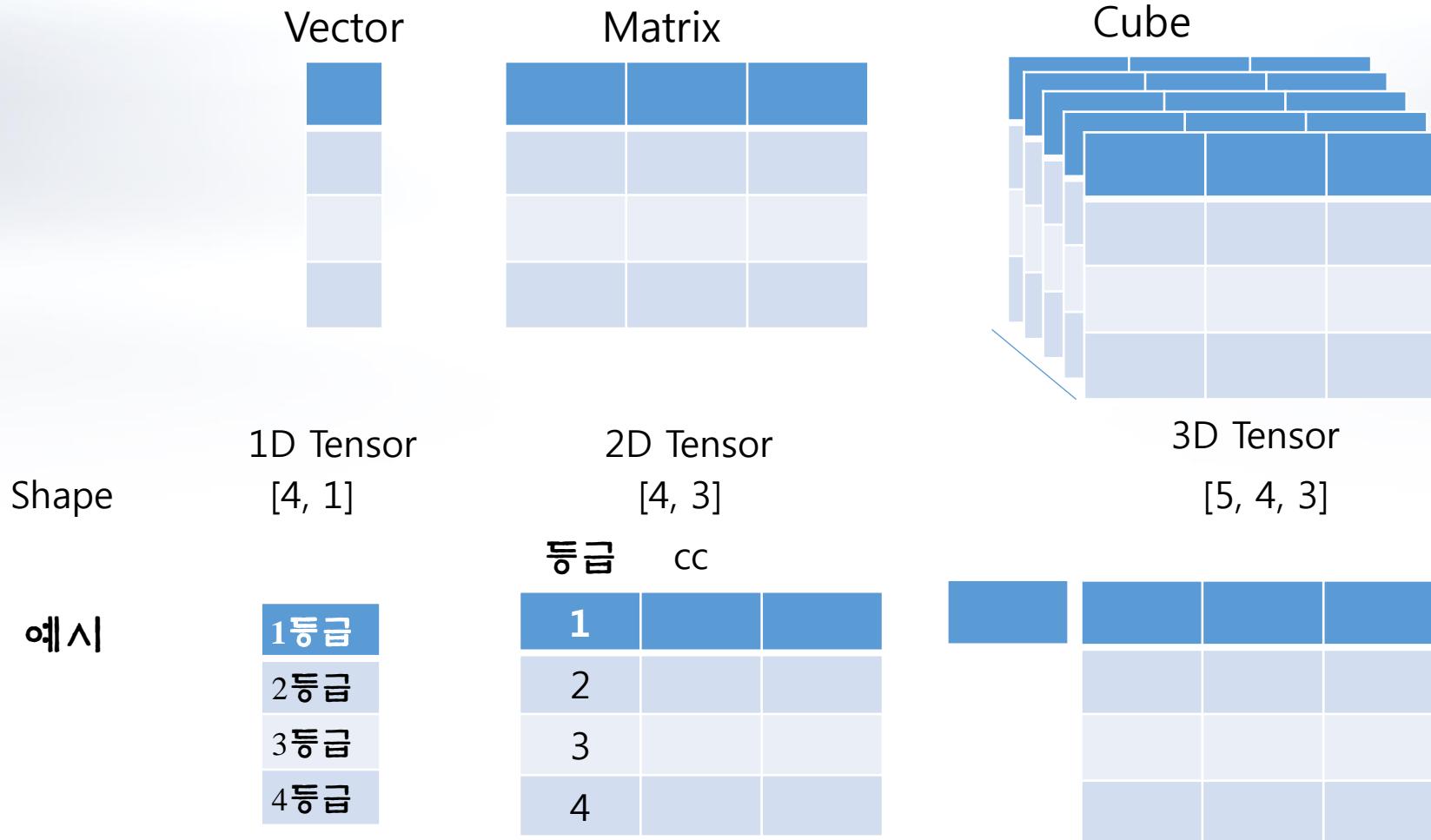
[source] <https://www.jamesserra.com/archive/2015/08/relational-databases-vs-non-relational-databases/>

3. 기본 지식과 기술(실습)

34

- Mathematics

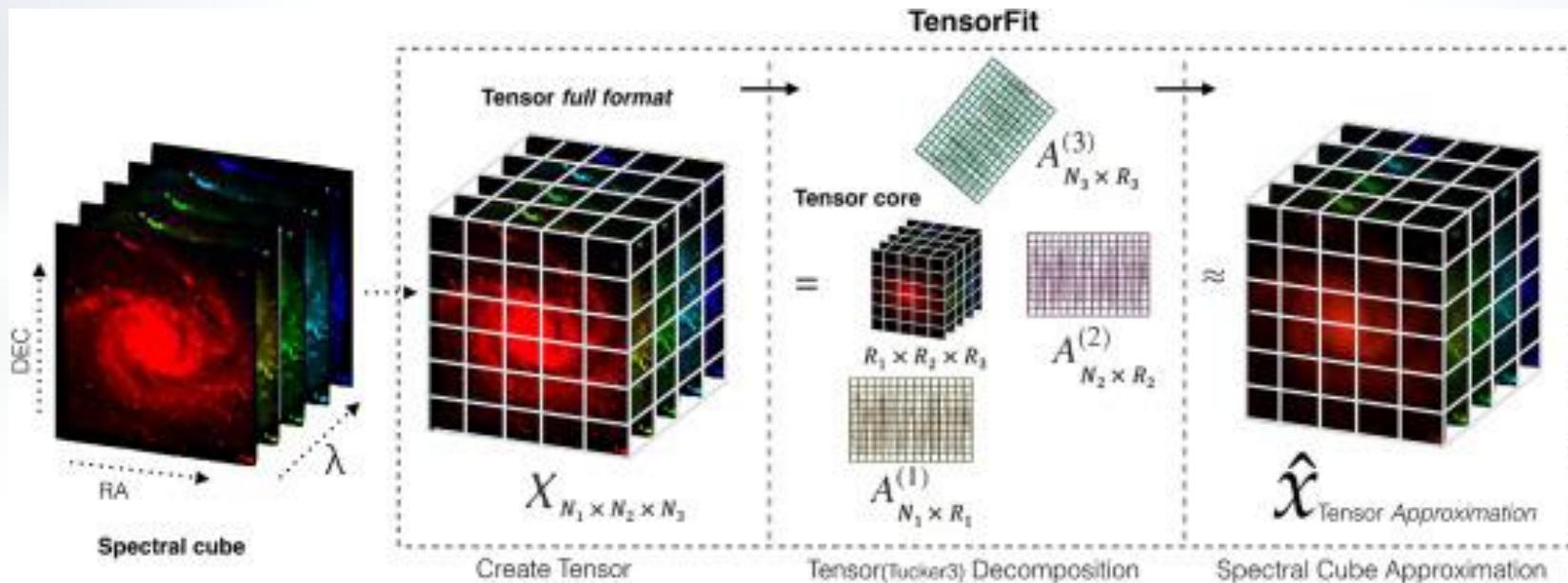
- 벡터, 행렬, 텐서



3. 기본 지식과 기술(실습)

35

- Mathematics
 - 텐서



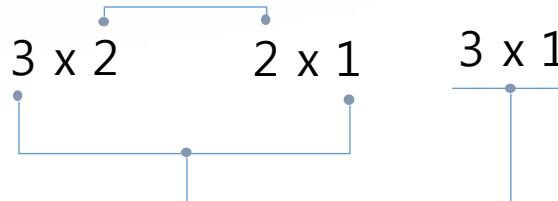
[source] <https://www.sciencedirect.com/science/article/pii/S221313371830043X>

3. 기본 지식과 기술(실습)

36

- Mathematics > 백터, 행렬, 텐서
 - Matrix Multiplication

$$\begin{matrix} 0 & 1 \\ 2 & 3 \\ 4 & 5 \end{matrix} \times \begin{matrix} 0.0 \\ 0.2 \end{matrix} = \begin{matrix} ? \\ ? \\ ? \end{matrix}$$



0.0	$0 * 0.0 + 1 * 0.2 = 0.2$
0.2	$2 * 0.0 + 3 * 0.2 = 0.6$
	$4 * 0.0 + 5 * 0.2 = 1.0$

3. 기본 지식과 기술(실습)

- Mathematics
 - 선형대수(Linear Algebra) or Try & Error

Prediction	Ground True	error
$\begin{bmatrix} \text{pred_label 1} \\ \text{pred_label 2} \\ \vdots \\ \text{pred_label n} \end{bmatrix}$	$\begin{bmatrix} \text{label 1} \\ \text{label 2} \\ \vdots \\ \text{label n} \end{bmatrix}$	$\begin{bmatrix} \text{error 1} \\ \text{error 2} \\ \vdots \\ \text{error n} \end{bmatrix}$
$n \times 1$	$n \times 1$	$= \text{cost} = \text{loss}$
$\begin{bmatrix} \text{feature 1} & 1 \\ \text{feature 2} & 1 \\ \vdots & \vdots \\ \vdots & \vdots \\ \text{feature n} & 1 \end{bmatrix}$	$\begin{bmatrix} \text{Weight} \\ \text{bias} \end{bmatrix}$	$=$
$n \times 2$	2×1	

$$\text{pred_label} = \text{Weight} * \text{feature} + \text{bias}$$

$$\text{Weight} = \frac{\sum_{i=0}^n (\text{feature} - E(\text{feature}))(\text{label} - E(\text{label}))}{\sum_{i=0}^n (\text{feature} - E(\text{feature}))}$$

3. 기본 지식과 기술(실습)

38

- Mathematics
 - 미분

$$Loss(weight) = \frac{1}{n} \sum_{i=0}^n (Hypothesis_{weight}(feature^{(i)}) - label^{(i)})^2$$

- $Loss(weight)$: Loss Function = Cost Function = MSE(Mean Squared Error)
- $feature^{(i)}$: i번째 관찰치의 input vector
- $weight$: feature의 계수값으로 정한 가중치
- $label^{(i)}$: i번째 관찰치의 실제 label (= target)
- $Hypothesis_{weight}(feature^{(i)})$: Weight parameter값을 i번째 관찰치에 학습 시켜 예측한 label 값

3. 기본 지식과 기술(실습)

39

- Mathematics
 - 미분

$$Loss(weight) = \frac{1}{n} \sum_{i=0}^n (Hypothesis_{weight}(feature^{(i)}) - label^{(i)})^2$$

$$\frac{\partial}{\partial \theta_j} \text{MSE}(\theta) = \frac{2}{m} \sum_{i=1}^m (\theta^T \cdot \mathbf{x}^{(i)} - y^{(i)}) x_j^{(i)}$$

3. 기본 지식과 기술(실습)

- Mathematics
 - Mathematical Function

Name	Plot	Equation	Derivative
Identity		$f(x) = x$	$f'(x) = 1$
Binary step		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ ? & \text{for } x = 0 \end{cases}$
Logistic (a.k.a Soft step)		$f(x) = \frac{1}{1 + e^{-x}}$	$f'(x) = f(x)(1 - f(x))$
Tanh		$f(x) = \tanh(x) = \frac{2}{1 + e^{-2x}} - 1$	$f'(x) = 1 - f(x)^2$
ArcTan		$f(x) = \tan^{-1}(x)$	$f'(x) = \frac{1}{x^2 + 1}$
Rectified Linear Unit (ReLU)		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Parameteric Rectified Linear Unit (PReLU) [2]		$f(x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Exponential Linear Unit (ELU) [3]		$f(x) = \begin{cases} \alpha(e^x - 1) & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
SoftPlus		$f(x) = \log_e(1 + e^x)$	$f'(x) = \frac{1}{1 + e^{-x}}$

tensorflow

- tf.nn.sigmoid
- tf.nn.relu
- tf.tanh
- tf.nn.softmax
- (and more)

3. 기본 지식과 기술(실습)

41

- Mathematics
 - Mathematical Function
 - Sigmoid function = Logistic function

```
#! /usr/bin/env python
"""
Program Name : func-sigmoid.py
"""

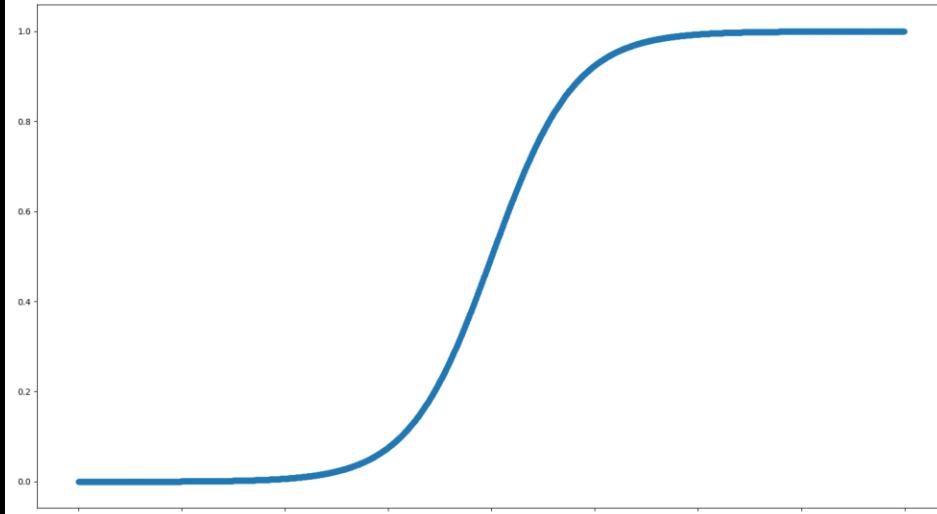
import numpy as np
def sigmoid(array):
    return 1 / (1 + np.exp(-array))

feature = np.arange(-10, 10, 0.01)
activatedFeature = sigmoid(array = feature)
print(activatedFeature)

import matplotlib.pyplot as plt
plt.scatter(feature, activatedFeature)
plt.show()
```

$$\text{Sigmoid}(\text{feature}) = \frac{1}{1 + e^{-\text{feature}}}$$

$e = \text{euler's number} = 2.718281828459045$

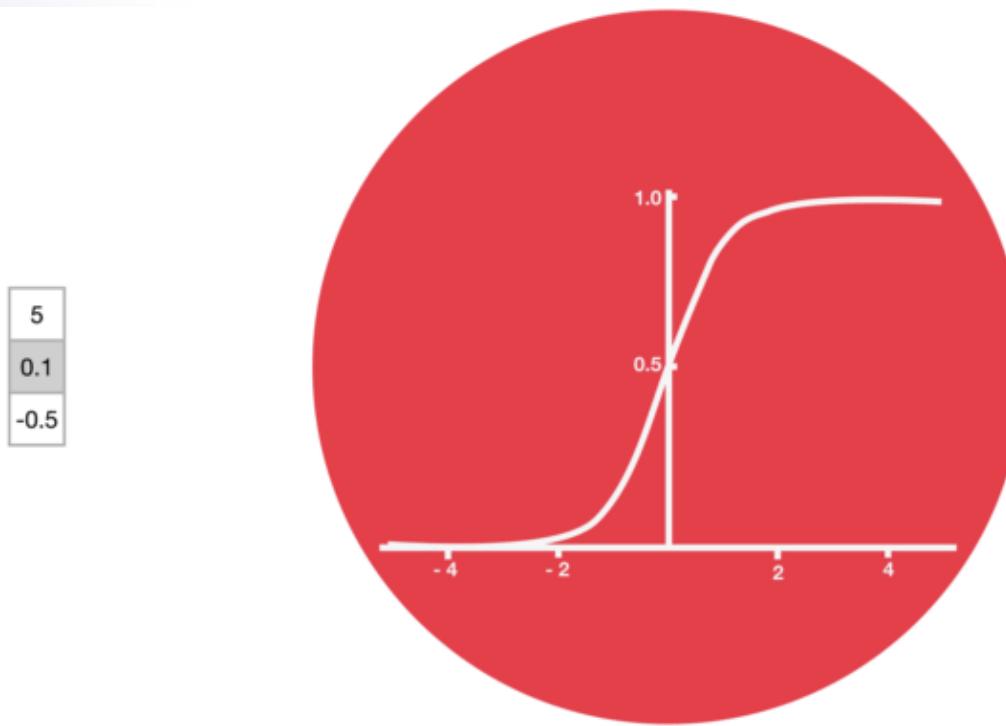


3. 기본 지식과 기술(실습)

42

- Mathematics
 - Mathematical Function
 - Sigmoid Function 의 효과

Activation!!!



3. 기본 지식과 기술(실습)

- Mathematics
 - Mathematical Function
 - Derivative of Sigmoid function = Logistic function

$$\begin{aligned}
 \frac{d}{dx} \sigma(x) &= \frac{d}{dx} \left(\frac{1}{1 + e^{-x}} \right) \\
 &= \frac{-(1 + e^{-x})'}{(1 + e^{-x})^2} \\
 &= \frac{e^{-x}}{(1 + e^{-x})^2} \\
 &= \left(\frac{1}{1 + e^{-x}} \right) \left(\frac{e^{-x}}{1 + e^{-x}} \right) \\
 &= \left(\frac{1}{1 + e^{-x}} \right) \left(\frac{1 + e^{-x}}{1 + e^{-x}} - \frac{1}{1 + e^{-x}} \right) \\
 &= \sigma(x) \left(\frac{1 + e^{-x}}{1 + e^{-x}} - \sigma(x) \right) \\
 &= \sigma(x) (1 - \sigma(x))
 \end{aligned}$$

$$\text{Sigmoid(feature)} = \frac{1}{1 + e^{-\text{feature}}}$$

$e = \text{euler's number} = 2.718281828459045$

