

A Venn diagram illustrating the relationship between Artificial Intelligence, Machine Learning, and Deep Learning. It consists of three nested rounded rectangles. The outermost rectangle is white and labeled '인공지능' (Artificial Intelligence). Inside it is a light gray rectangle labeled '머신러닝' (Machine Learning). Inside the light gray rectangle is a darker gray rectangle labeled '딥러닝' (Deep Learning). This visualizes that Deep Learning is a subset of Machine Learning, which is a subset of Artificial Intelligence.

인공지능

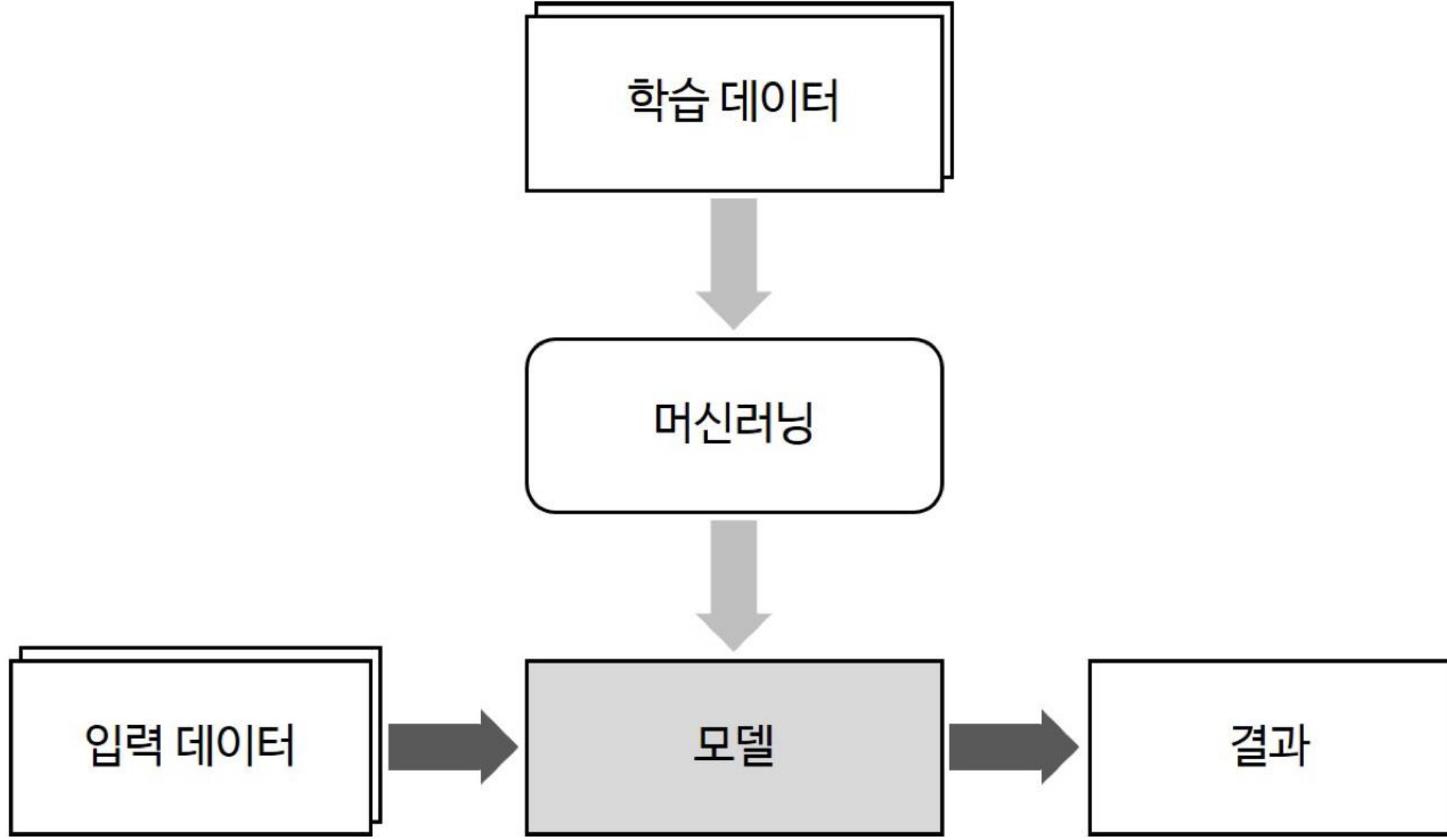
머신러닝

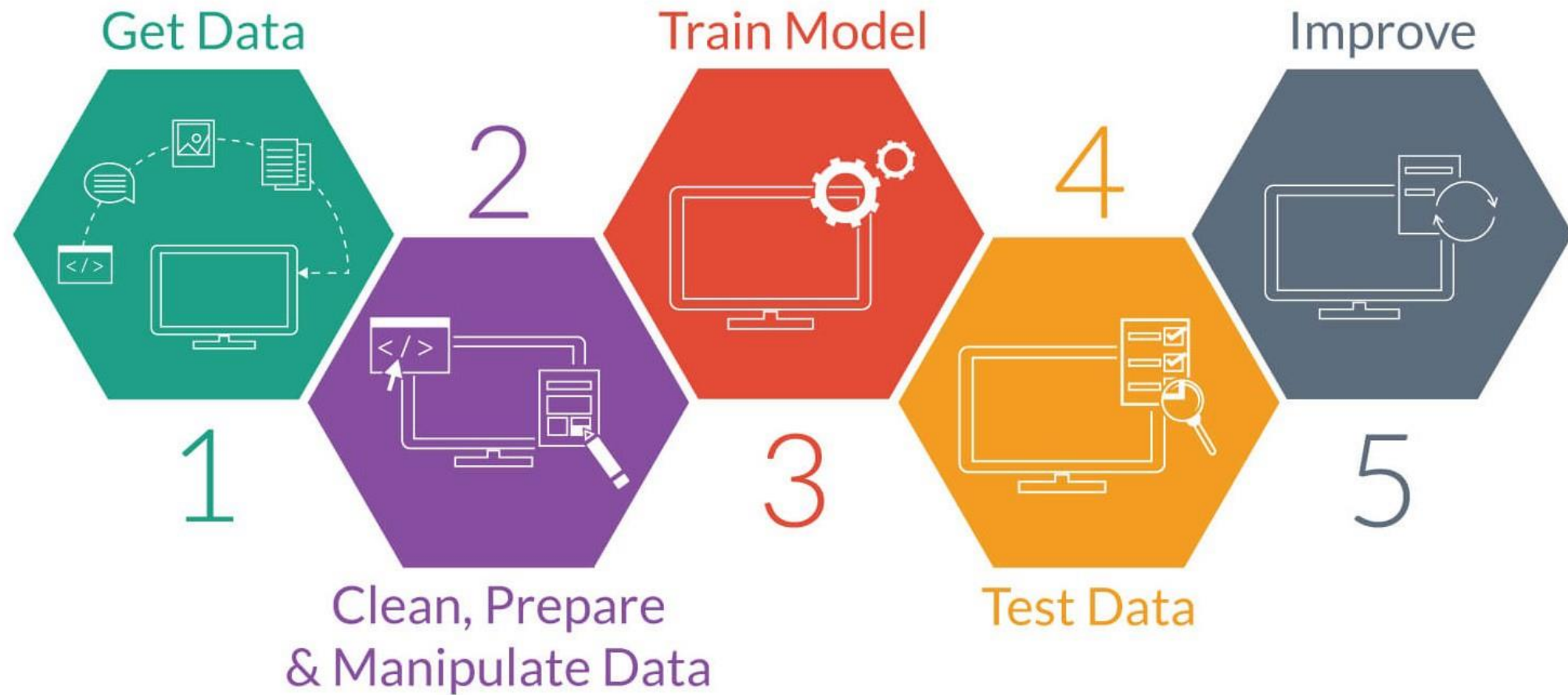
딥러닝

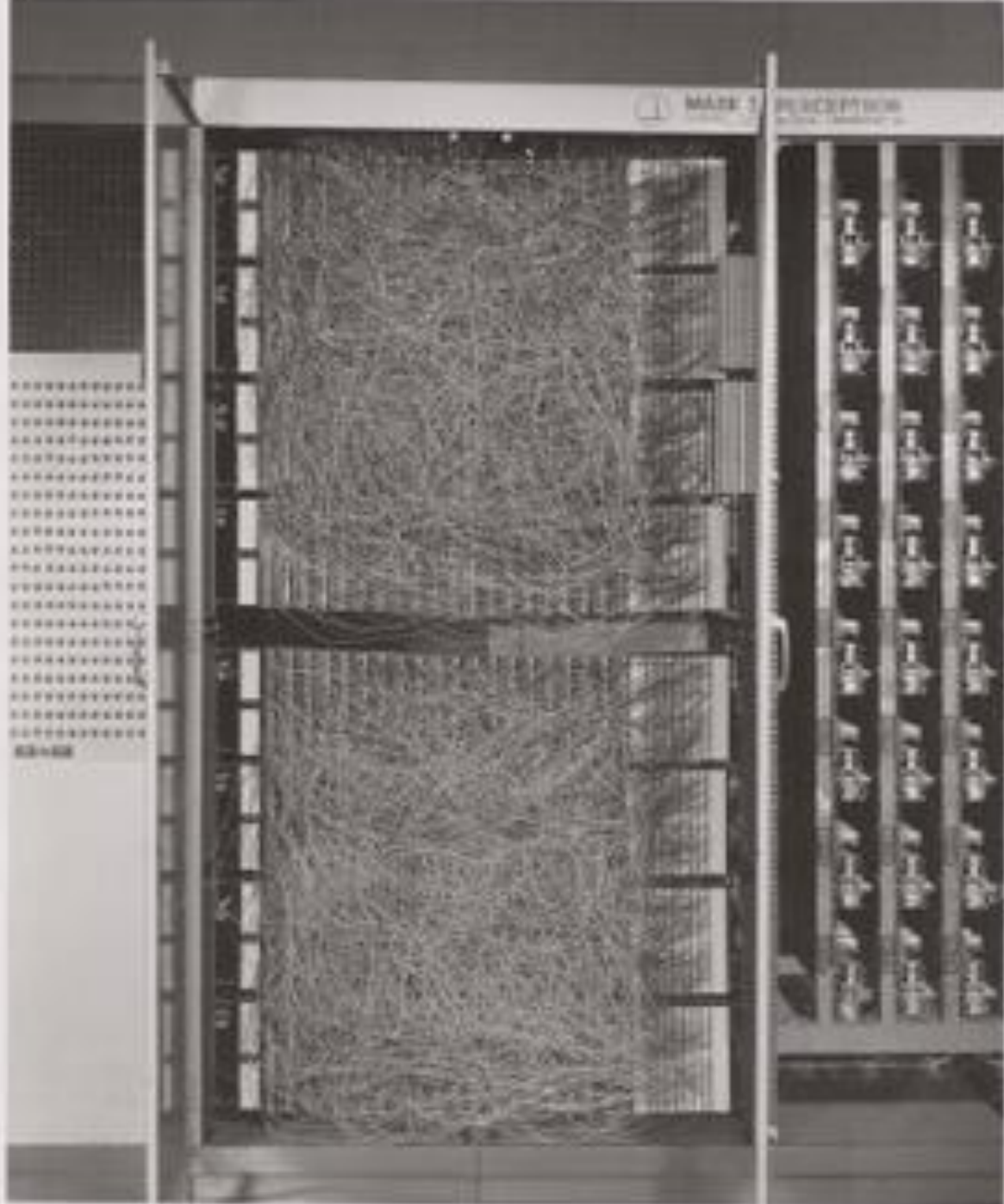


GARBAGE IN | GARBAGE OUT

- 머신러닝: **데이터**에 기반한 **모델링** 기법
- 데이터의 영향을 매우 크게 받는다







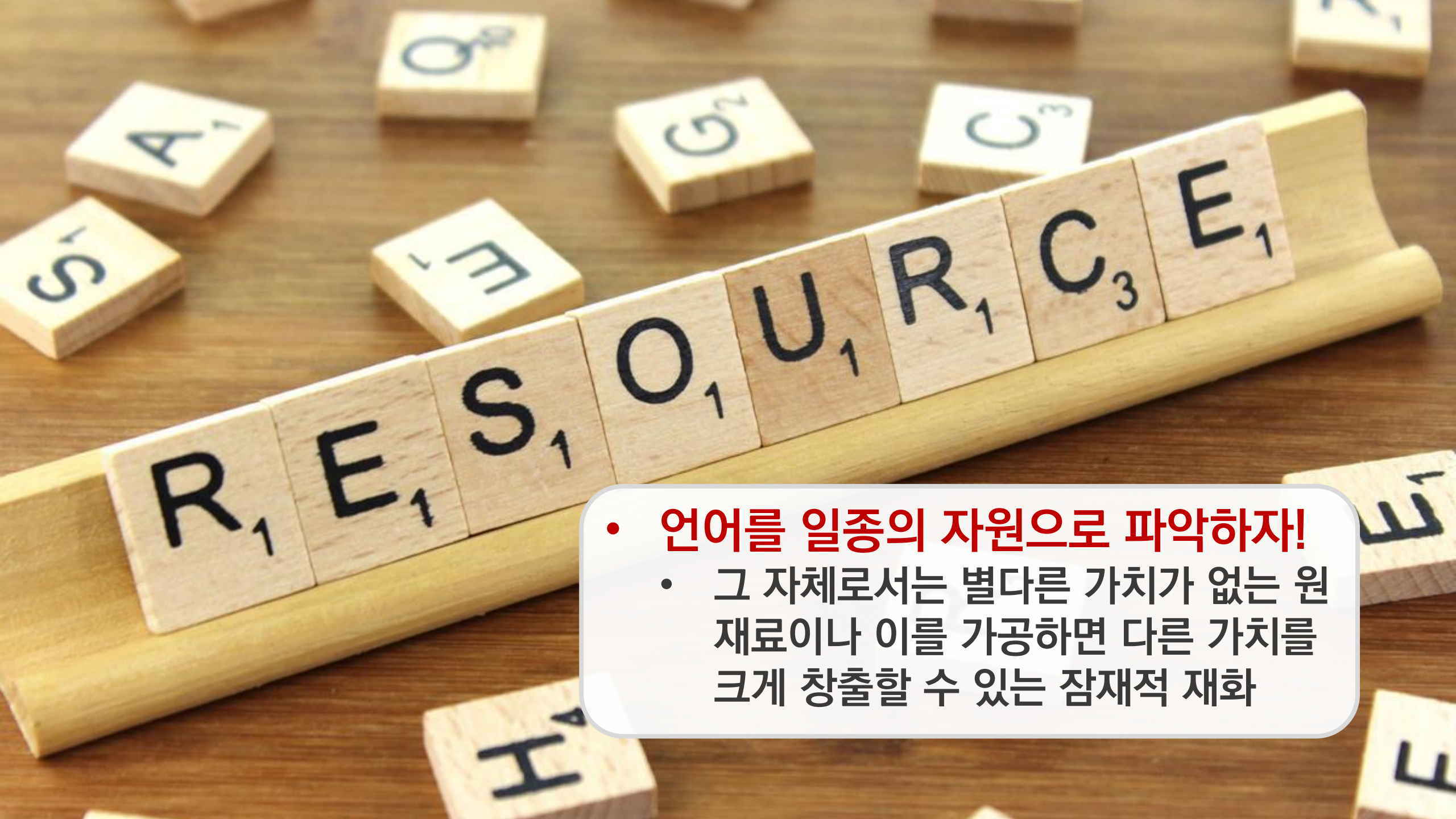
Perceptron

알고리즘의 정비 기계적 뒷받침 데이터의 확보

충분히 뛰어난 모델
충분히 좋은 컴퓨터
충분히 큰 자원

빅
데이터

자원으로서의 언어



- 언어를 일종의 자원으로 파악하자!
 - 그 자체로서는 별다른 가치가 없는 원재료이나 이를 가공하면 다른 가치를 크게 창출할 수 있는 잠재적 재화

과적합: 過猶不及

高校英語の 礎石



성문 基本英語

宋成文 著



成文出版社

속 어

come true 실현되다
come up to ~에 이르다
come up with ~을 생각해내다
compare A to B = A를 B에 비유하다
be composed of ~으로 구성되다
be concerned about ~을 걱정하다
be concerned with ~와 관계가 있다
consist in ~에 (놓여) 있다
consist of ~으로 구성되다
correspond to ~에 일치하다
correspond with ~와 서신 교환을 하다
count for much 매우 중요하다

- (1) Everything he predicted has come true
- (2) His ability did not come up to me
- (3) An inventor is someone who usually comes up with ideas
- (4) Books are often compared to friends
- (5) The committee is composed of 10 members
- (6) I was very concerned about my mother
- (7) This story is concerned with a Russian
- (8) My pleasure consists mainly in helping others
- (9) The human body consists of billions of cells
- (10) The article doesn't correspond to the facts
- (11) I correspond with my friend in America
- (12) I soon realized that my experience counted for much
- (13) Knowledge without common sense counts for much to you to finish it

새 교과서 마스터中學數學

1

金教昌 著

圖書出版

書筆社

새 교육과정에 의한

마스터 中學數學

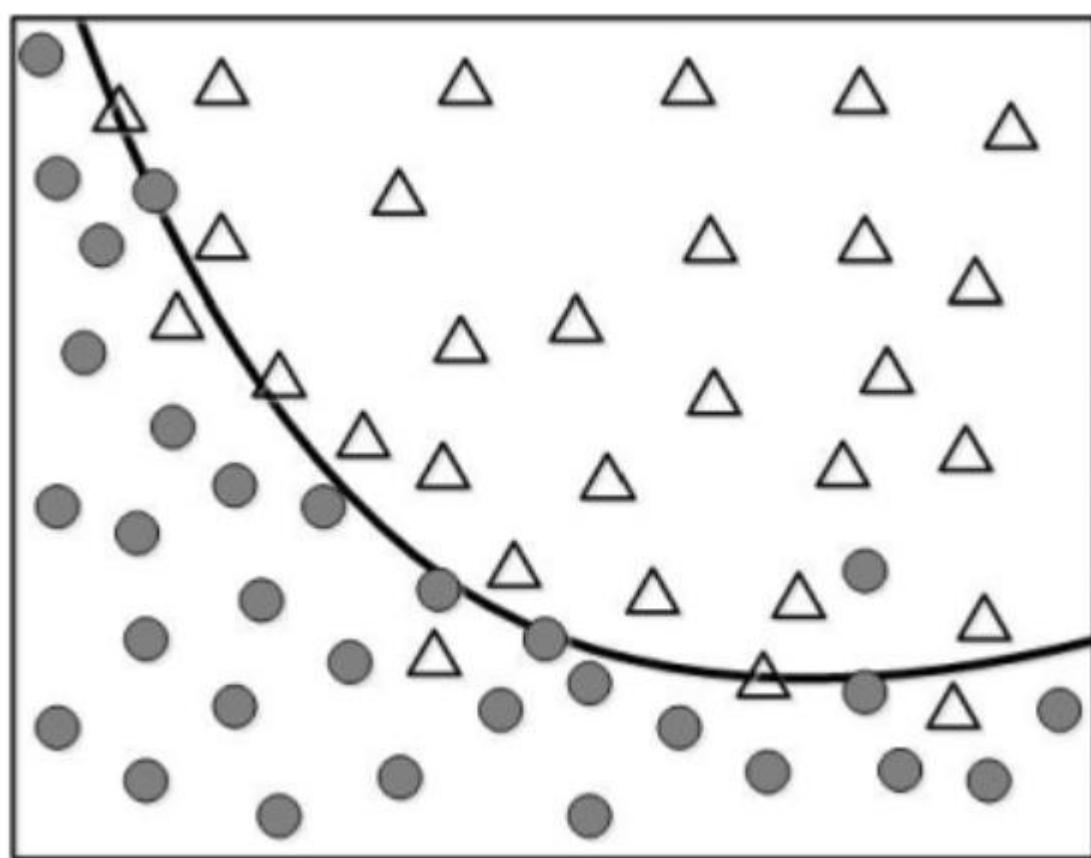
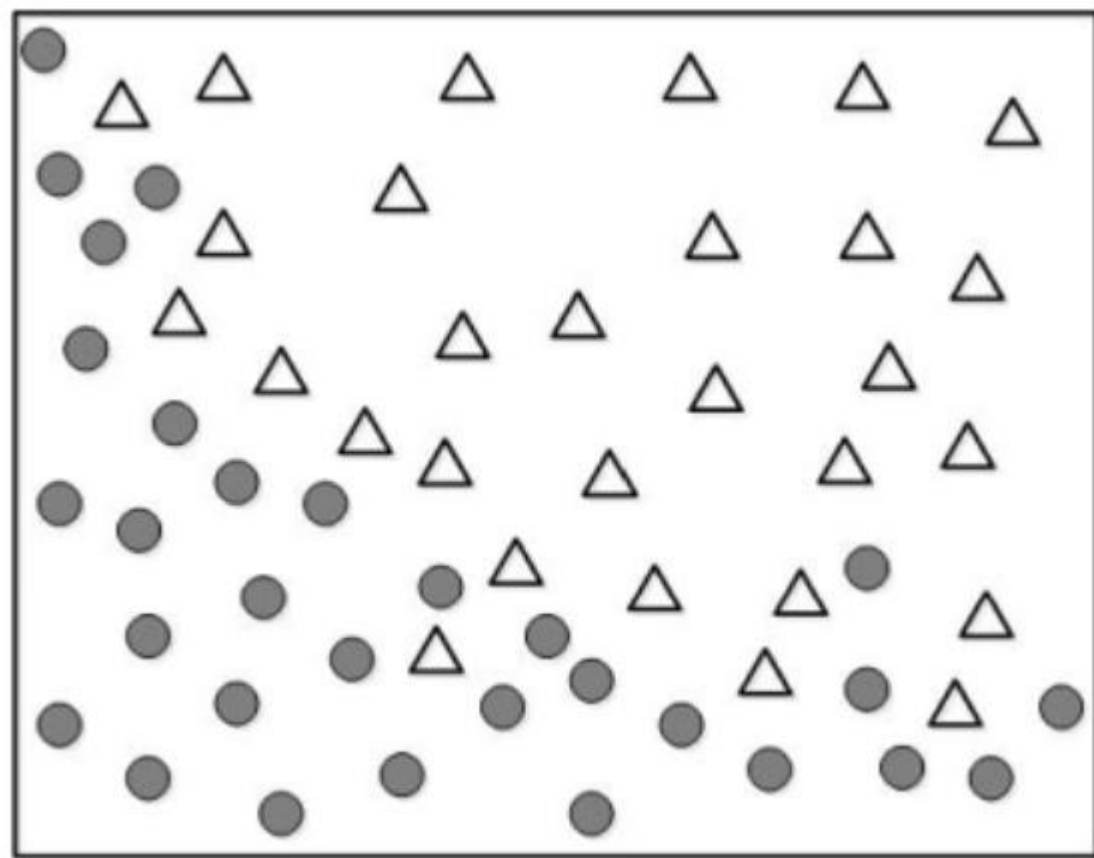


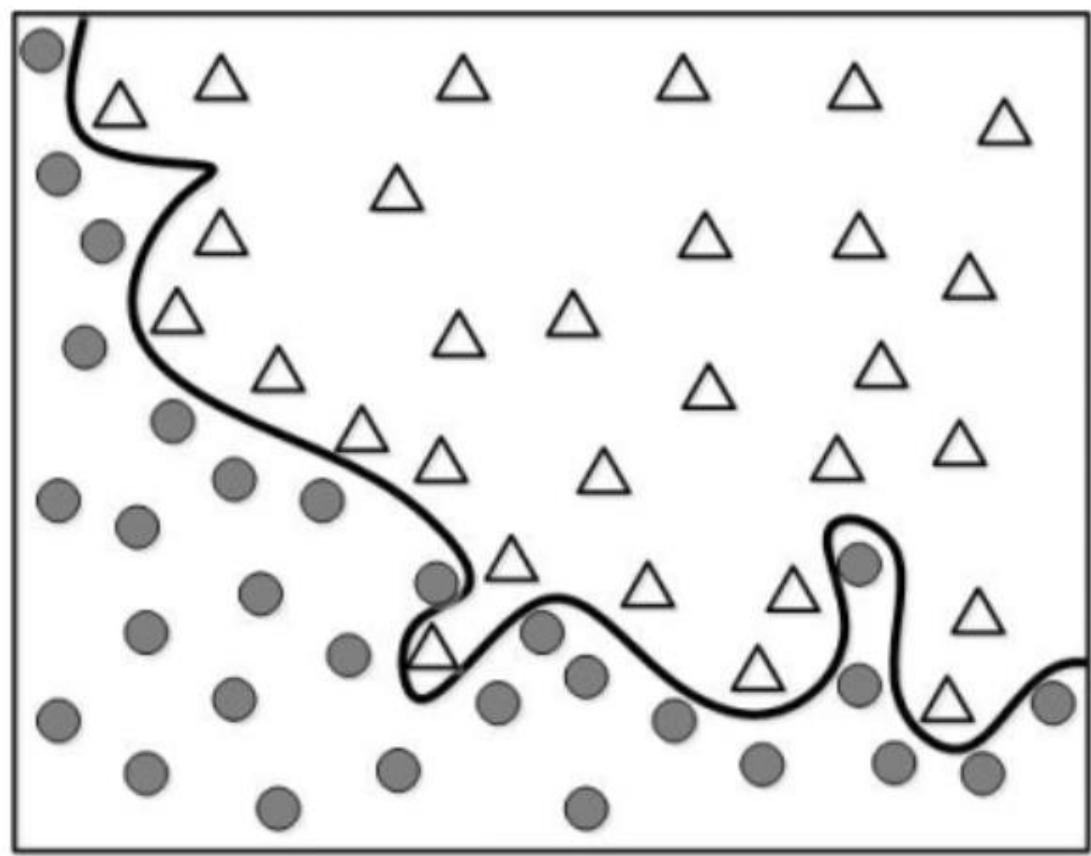
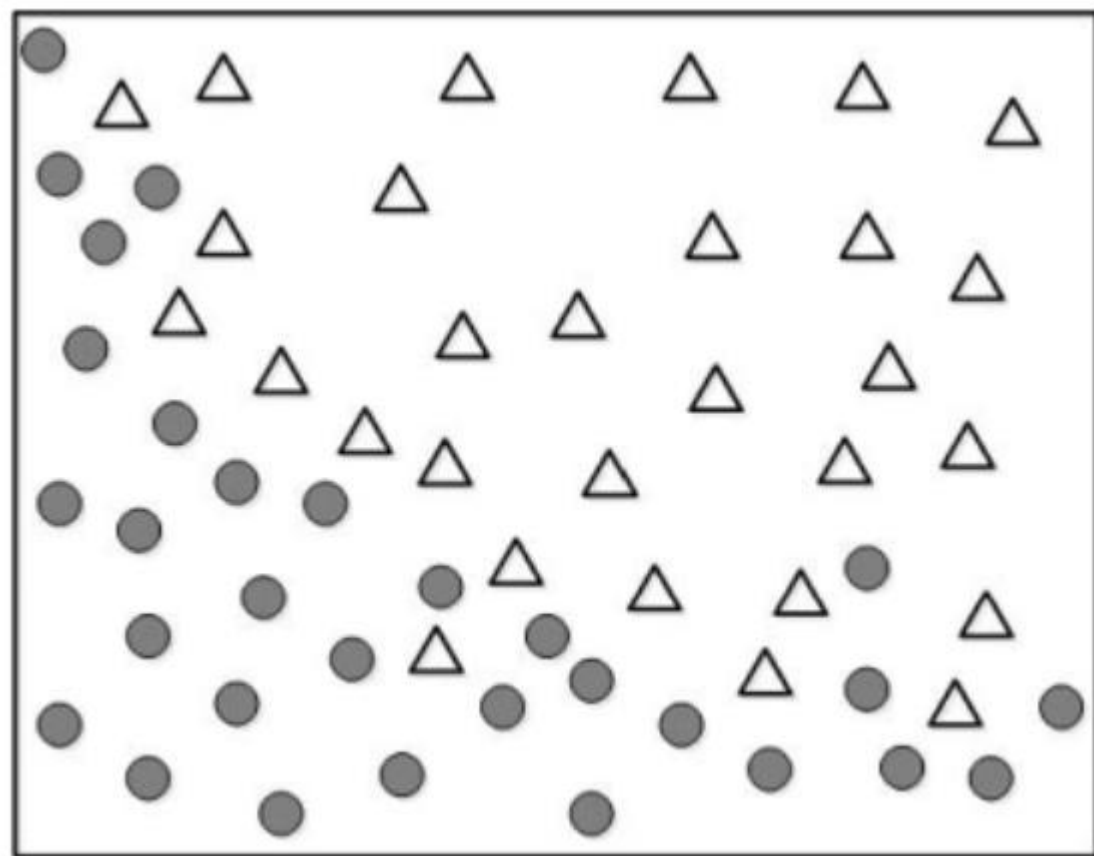
金教昌 著

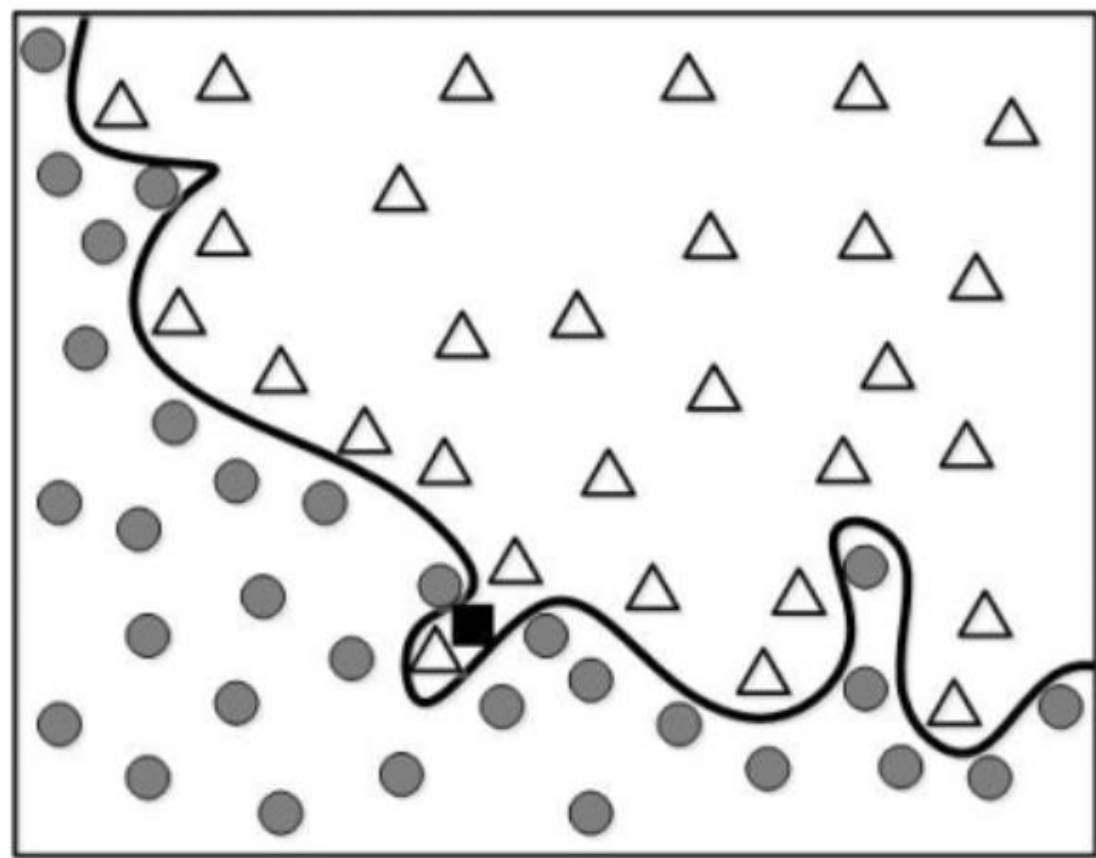
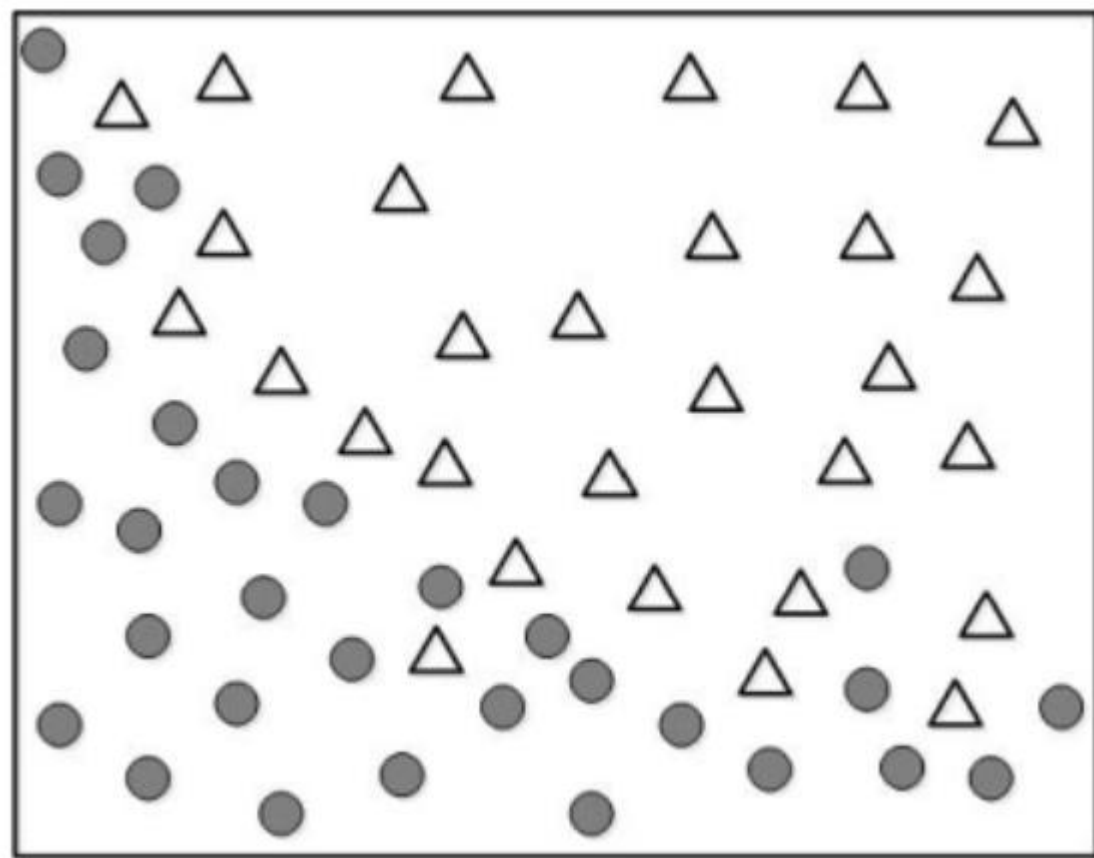


1

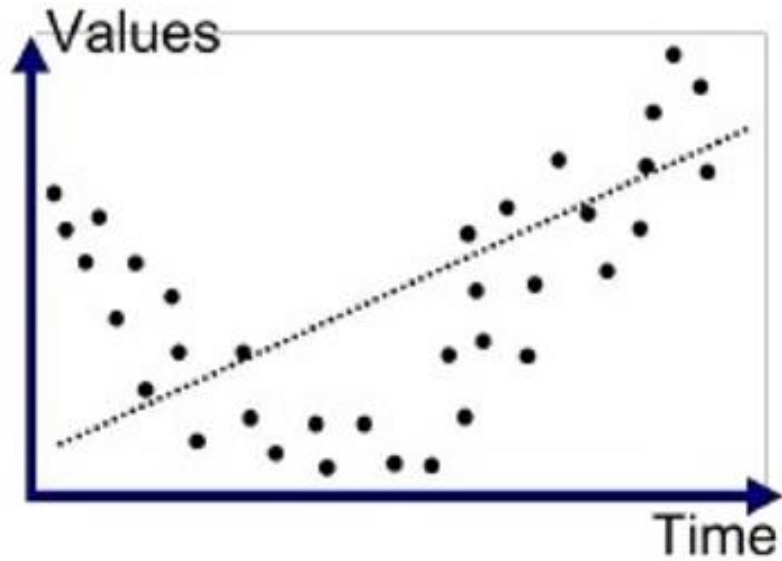
圖書出版 書筆社



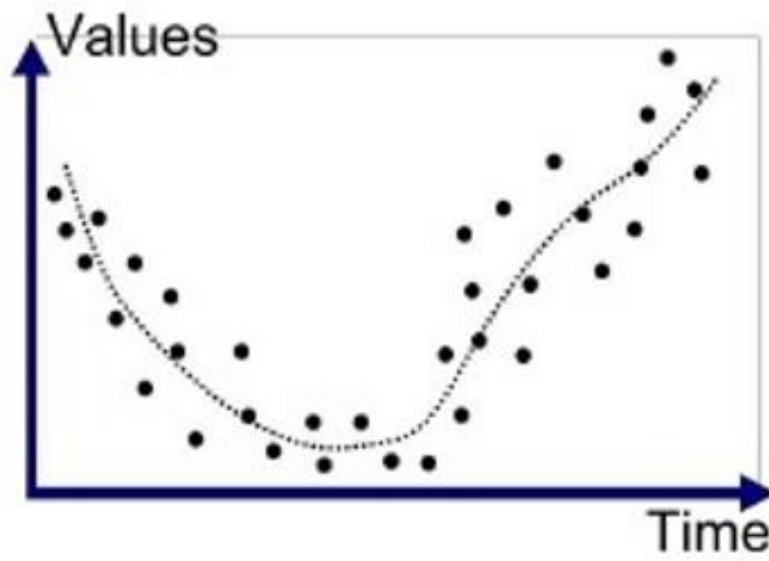




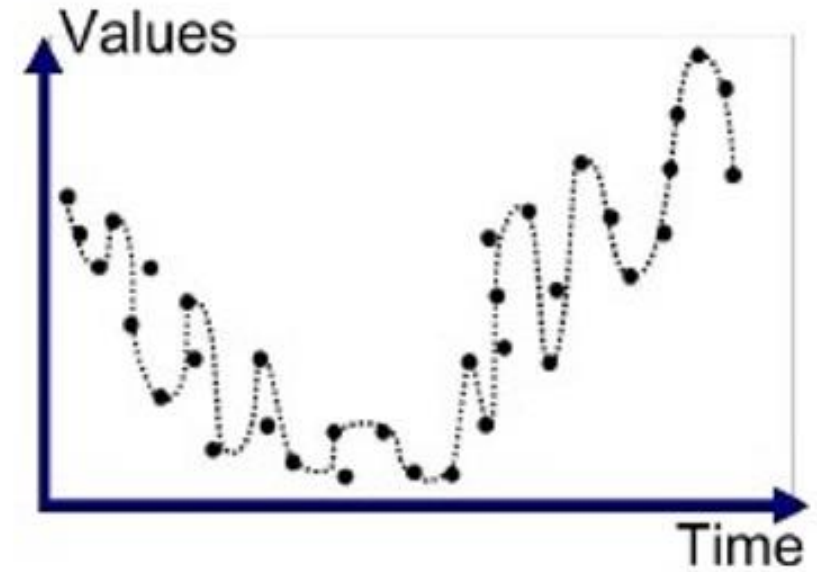
Training and Validation: What is Overfitting?



Underfitted



Good Fit/Robust



Overfitted

과적합과 싸우는 방법

- 정칙화(regularization)
 - 모델의 형태를 최대한 단순하게 만들려는 수치해석적인 기법
 - 오컴의 면도날
- 검증(validation)
 - 학습 데이터의 일부를 따로 떼어 내서
학습에는 사용하지 않고
모델의 성능 검증용으로 사용하는 기법

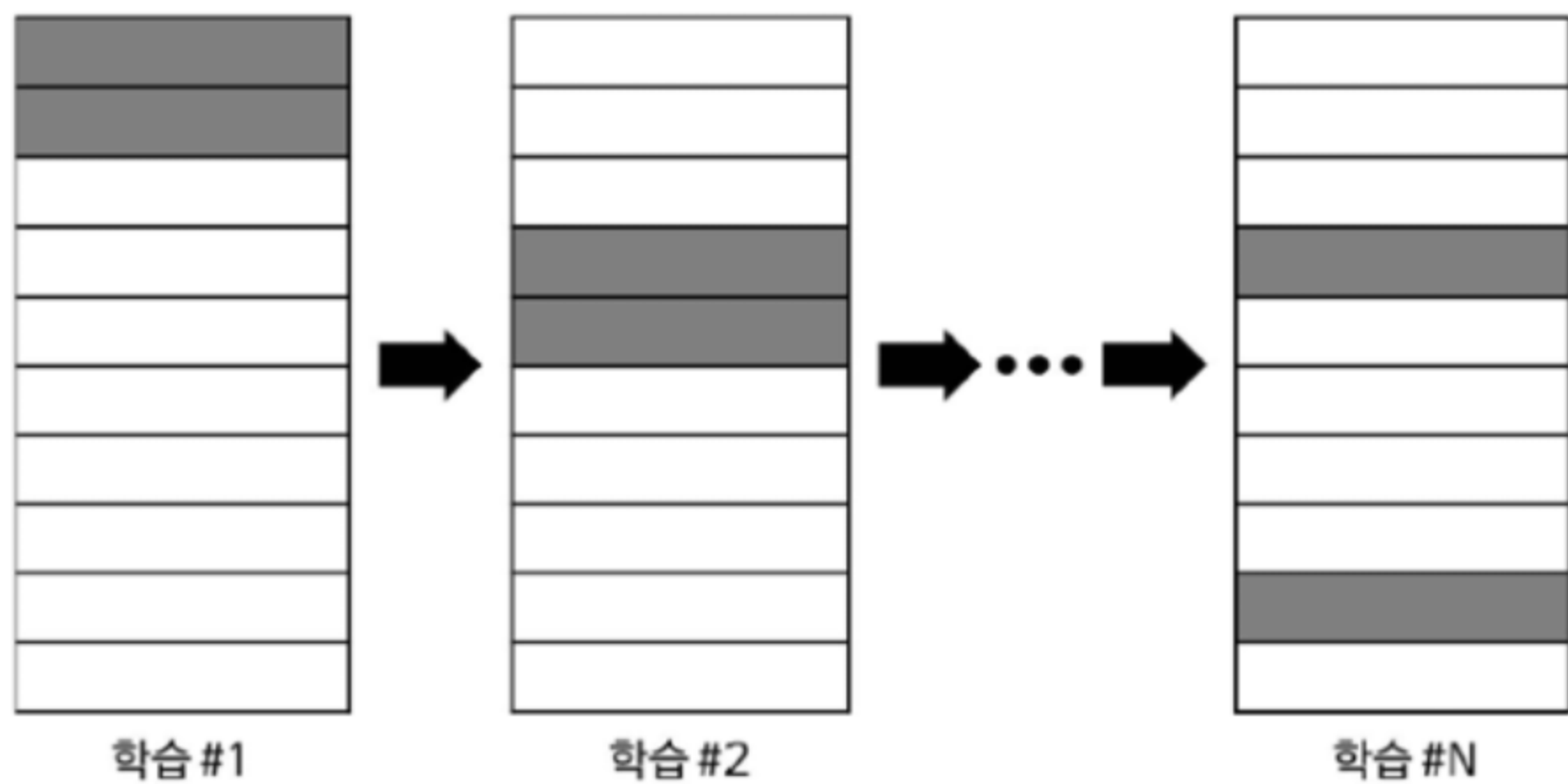


학습 데이터



학습 데이터

검증 데이터



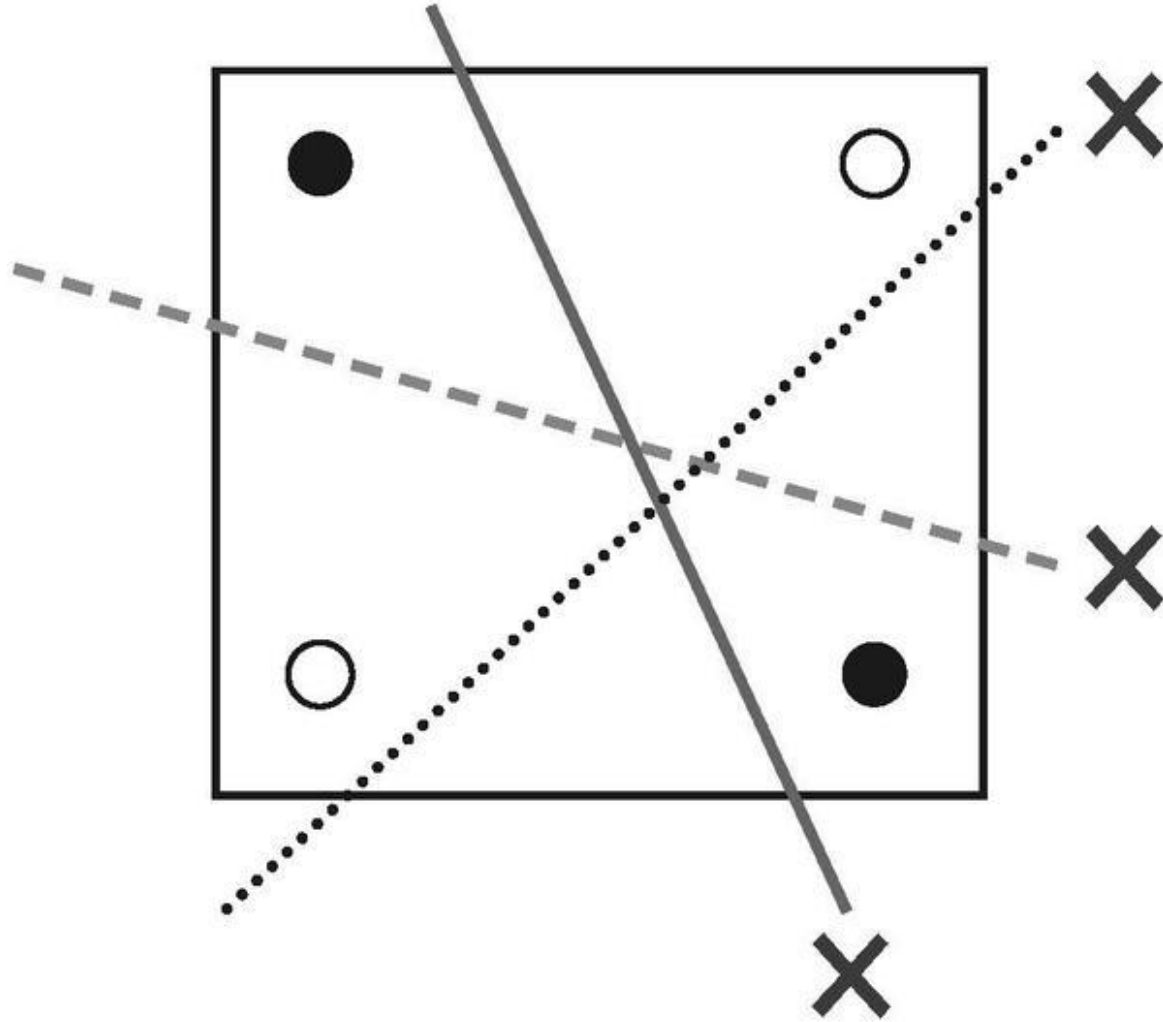
기계 학습의 종류

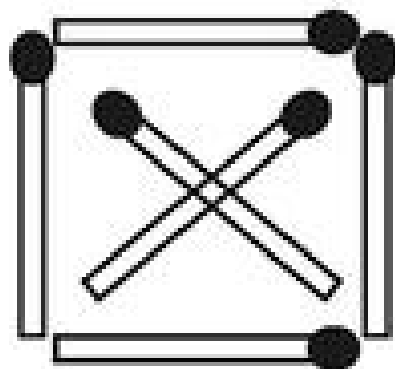
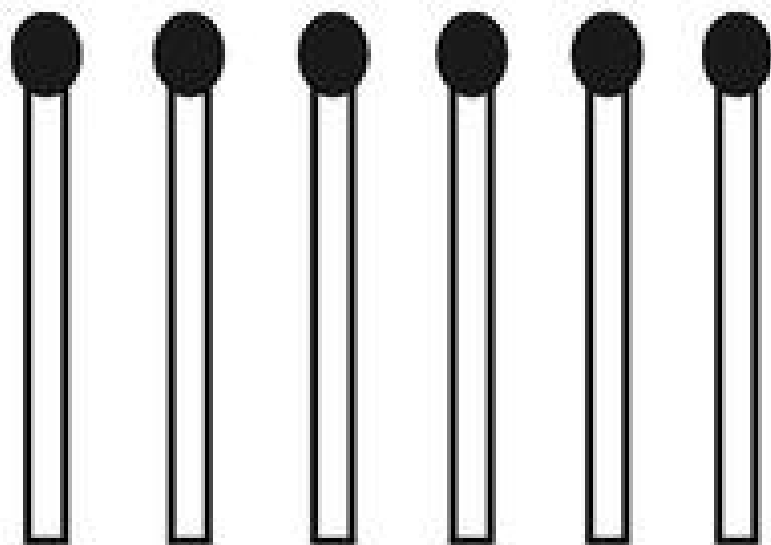
- 지도학습(supervised learning): {입력, 정답}
 - 분류(classification): 판별
 - binary, n-ary
 - 회귀(regression): 추세
- 비지도학습(unsupervised learning): {입력}
 - 분류
 - 군집화(clustering)
- 강화학습(reinforcement learning): {입력, 출력, 점수}

Deep Learning:

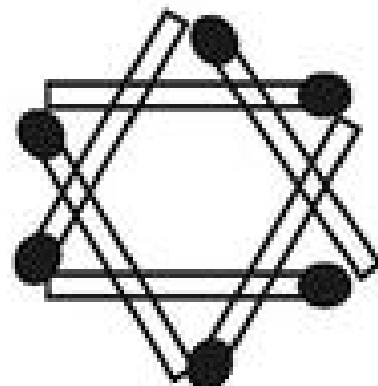
심층이 무엇을 말하는가?

복잡한 문제를 풀어야 하는데 단층 퍼셉트론은 단순한 XOR (exclusive-OR) 문제조차 풀기 어려워





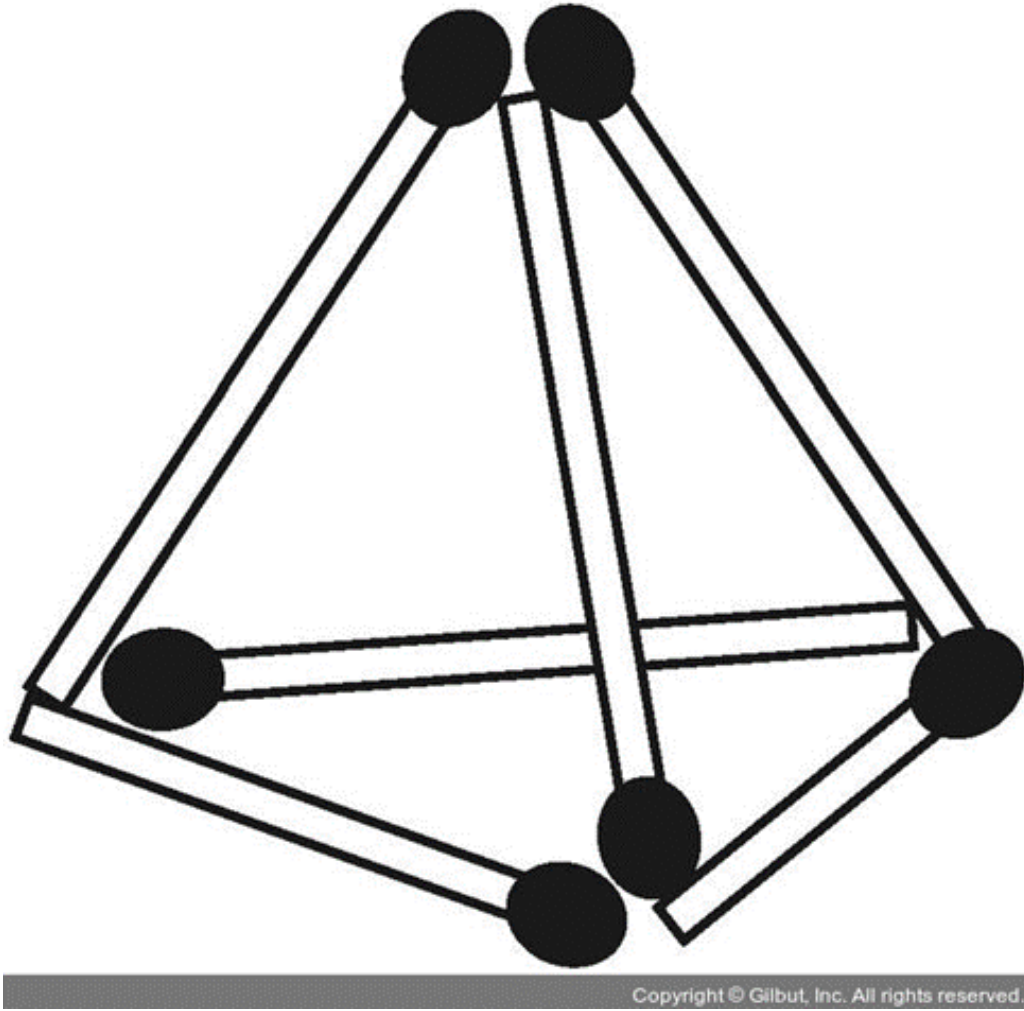
X



X

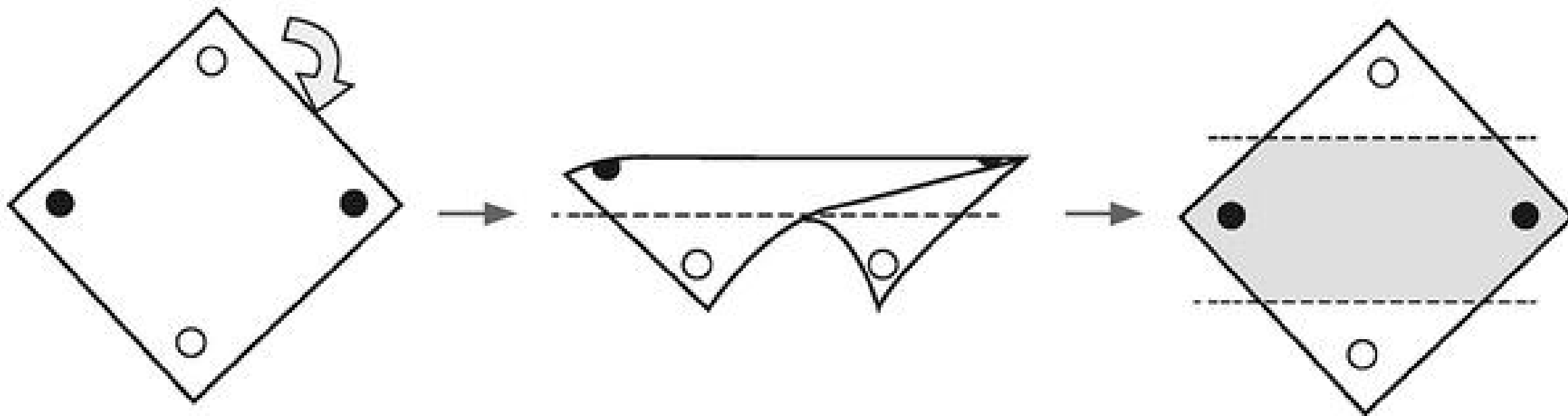
Copyright © Gilbut, Inc. All rights reserved.

성냥개비 여섯 개로 정삼각형 네 개를 만들 수 있는가?

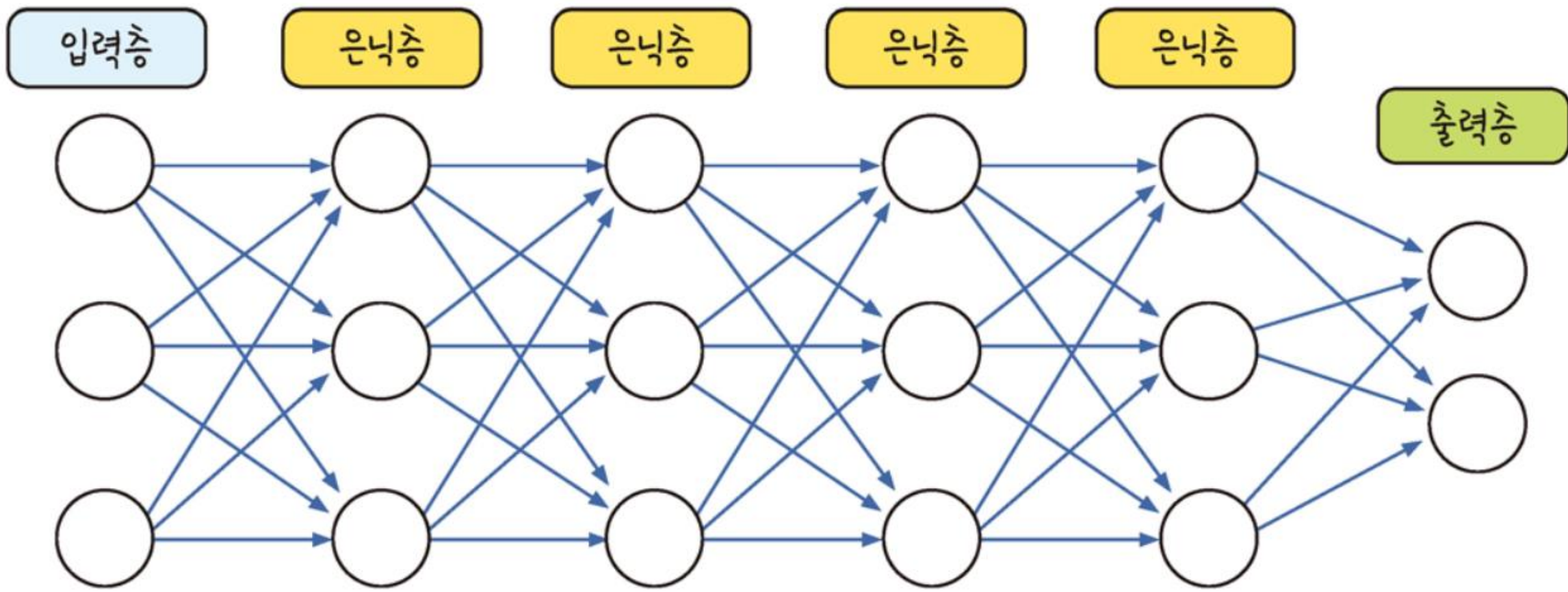


2차원에서만 문제를 해결하려는 고정관념을 깨자!

발상의 전환: 2차원에서 3차원으로



Copyright © Gilbut, Inc. All rights reserved.



Processor

CPU / GPU





OS



AisleRiot Solit...



Aktualisierung...



Amazon



Anwendungen ...



Aufgaben



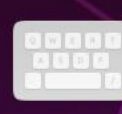
Cheese



Dateien



Dokumentenb...



Eingabemetho...



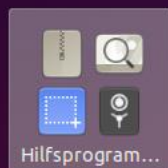
Einstellungen



Energienstatisti...



Firefox Web B...



Hilfsprogram...



Kalender



LibreOffice Calc



LibreOffice Dr...



LibreOffice Im...



LibreOffice W...



Livepatch



Mahjongg



Minen



Passwörter un...



Remmina



Rhythmbox

Häufig

Alle

바로 할 수 있는가?:
gpu, ubuntu

학교에서 사주는가?:
인문대가 그런 게 왜 필요해???

장착만 하면 되는가?:

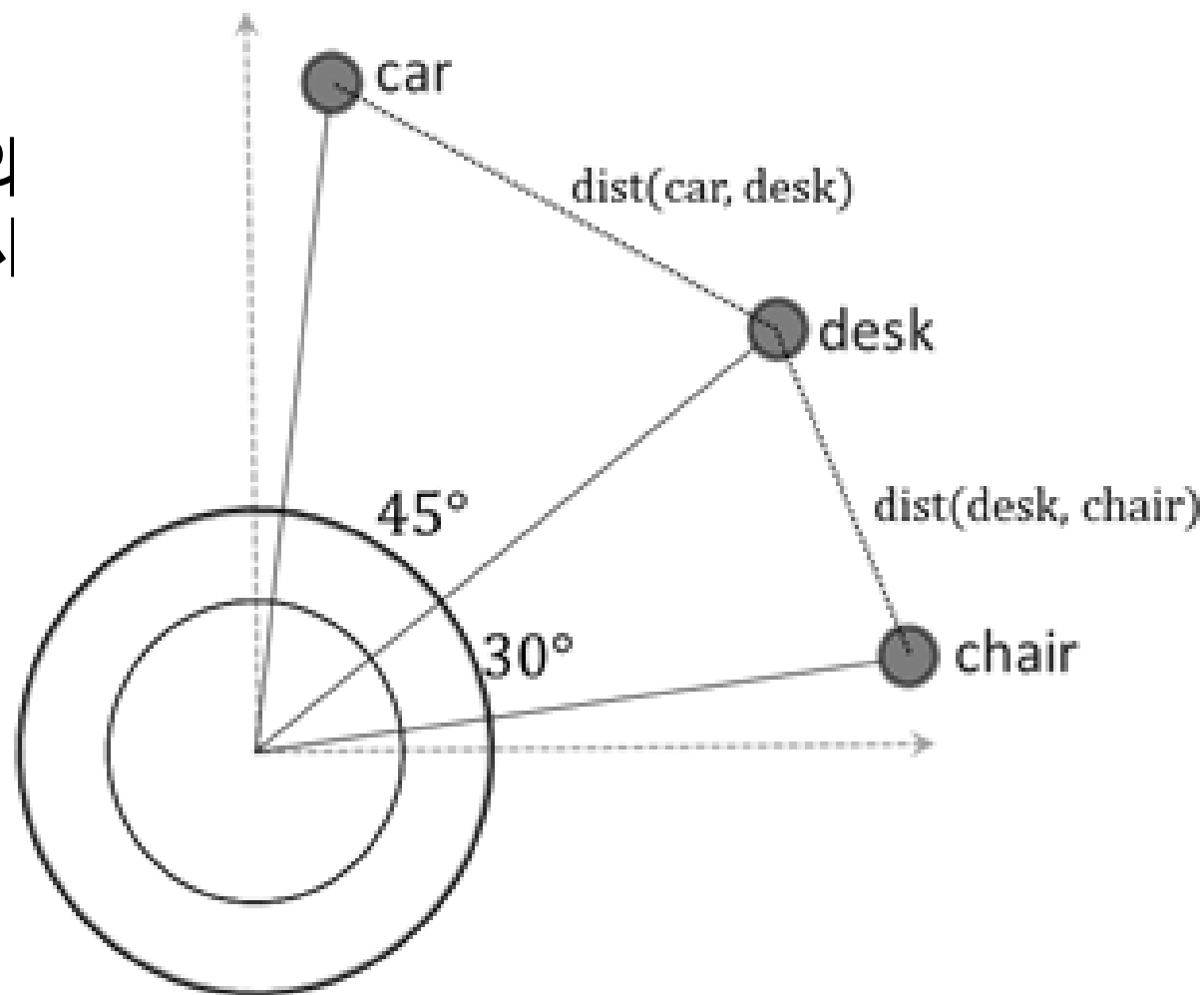
cpu, memory, power, cooler

Google Colab

단어/문장 관련도

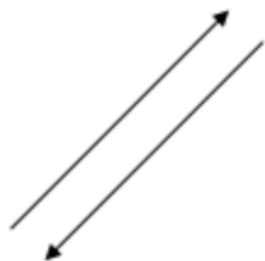
- 2차원(평면 공간) 안에 단어들의 위치를 찍을 수만 있다면, 단어들의 관련성, 어떤 단어들이 가깝고 먼지 알 수 있다.

- $\cos 45 = 0.5253\dots$
 $\cos 30 = 0.1543\dots$



코사인 유사도

- 일상 언어 속의 각(角): 관점, 형세, ...
 - 각이 나온다!, **각: 망신^각, 재수강^각, 숲속^각
- 두 벡터 간의 코사인각를 이용하여 구할 수 있는 두 벡터의 유사도.
 - 단어를 임베딩하여 수치화하였으면, 이러한 표현 방법에 대해서 코사인 값을 이용하여 유사도를 측정할 수 있다.



코사인 유사도 : -1



코사인 유사도 : 0



코사인 유사도 : 1

위대한 수령 김일성동지와
위대한 령도자 김정일동지
혁명사상만세!

로동신문

조선로동당 중앙위원회 기관지

제182호 [주체 제26421호] 주체 108 (2019)년 7월 1일 (월요일)

당의 령도따라 내 나라,
내 조국을 더욱 부강하게 하기
위해 힘차게 일해나가자!

경애하는 최고령도자 김정은동지께서 도널드 트럼프 미합중국 대통령과 판문점에서 역사적인 상봉을 하시였다

조선로동당 위원장이시며 조선민주주의인민공화국 국무위원회 위원장이시며 조선민주주의인민공화국 주체 최고사령관이신 우리 당과 국가, 주체의 최고령도자 김정은동지께서 6월 30일 오후 도널드 트럼프 미합중국 대통령의 세리에 마르 판문점에서 역사적인 상봉을 하시였다.

경애하는 최고령도자동지께서는 트럼프대통령이 6월 29일부터 30일까지 남조선을 방문하는 기회에 비무장지대에서 김정은 국무위원장과

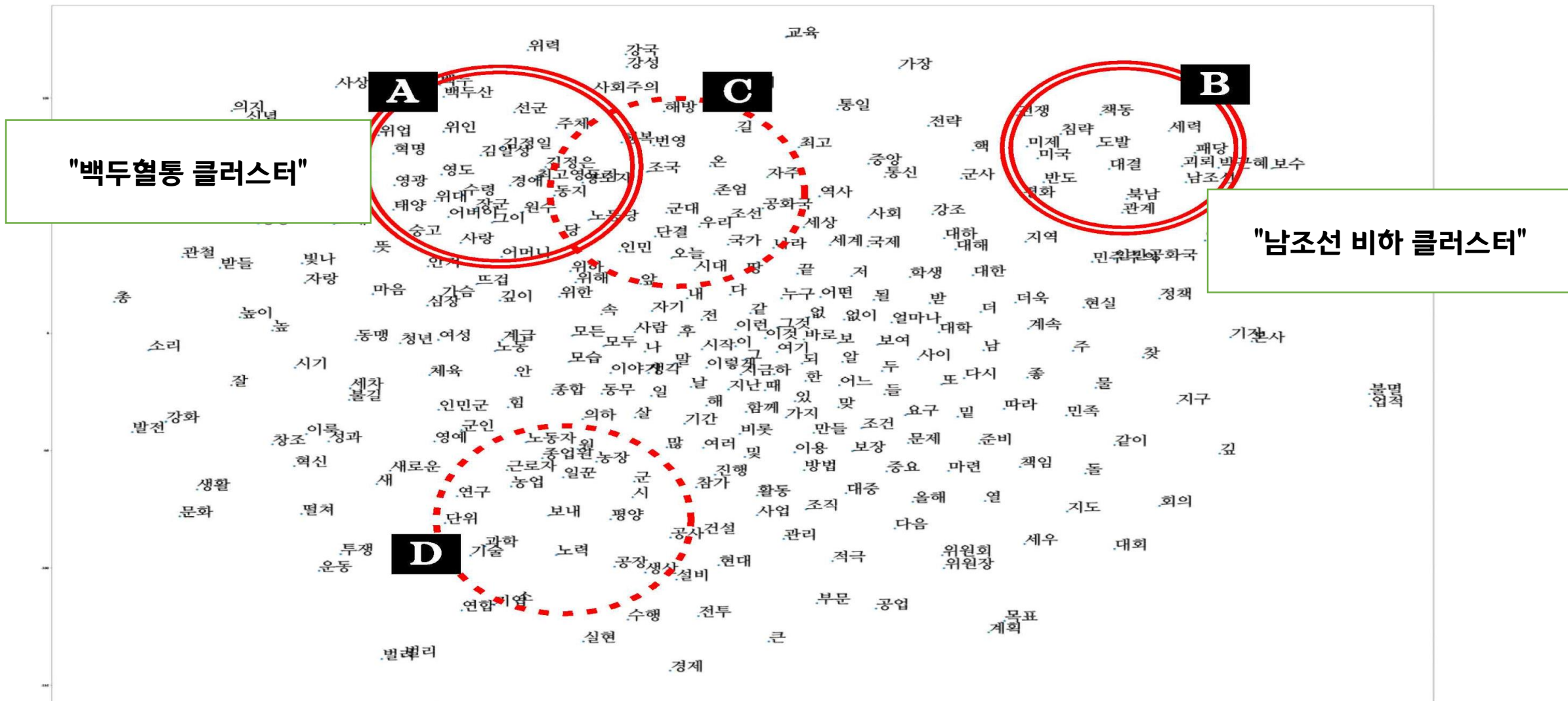


만나고싶다는 의사를 전해온데 대하여 수락하시고 판문점 남북지역에 나가셔서 트럼프대통령과 전격적으로 상봉하시였다.

트럼프대통령의 판문점상봉초청의사가 세상에 전해진 순간부터 국적인 사면이 이루어진 순간까지 하루 남짓한 시간동안 온 지구촌의 눈과 귀가 또다시 조선반도에로 집중되고 판문점에서의 초미수뇌상봉소식에 대한 관심과 기대가 온 행성을 뜨겁게 달구며 적절과 흥분으로 열광하였다.

2면으로 계속

정유남, 왕규현, 송상헌 (2020) 딥러닝을 이용한 노동신문의 이념적 어휘 연구



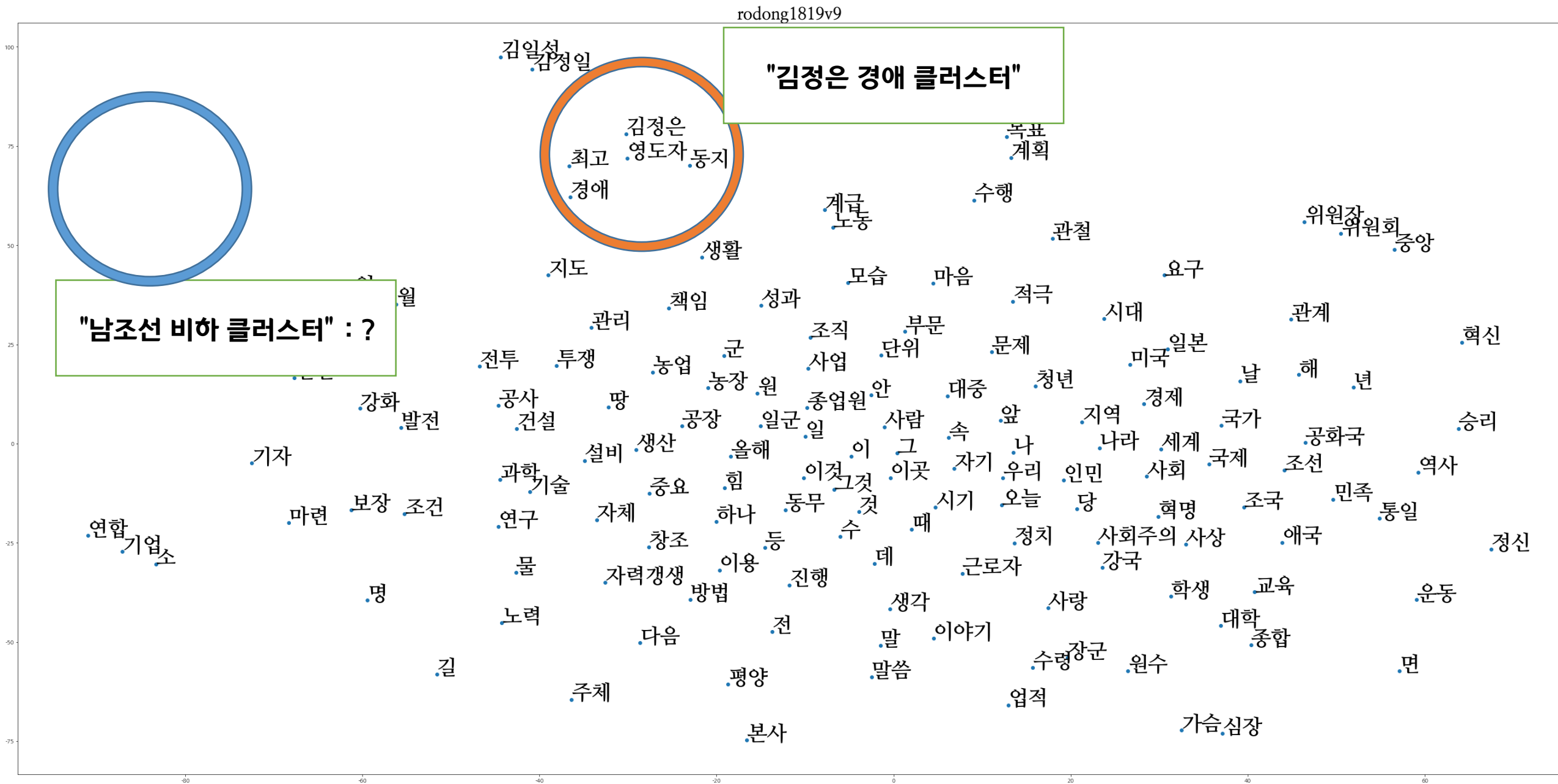
<그림 7> t-SNE 시각화(노동신문 5년치): 군집과 핵심 키워드

"김정은 경애 클러스터"

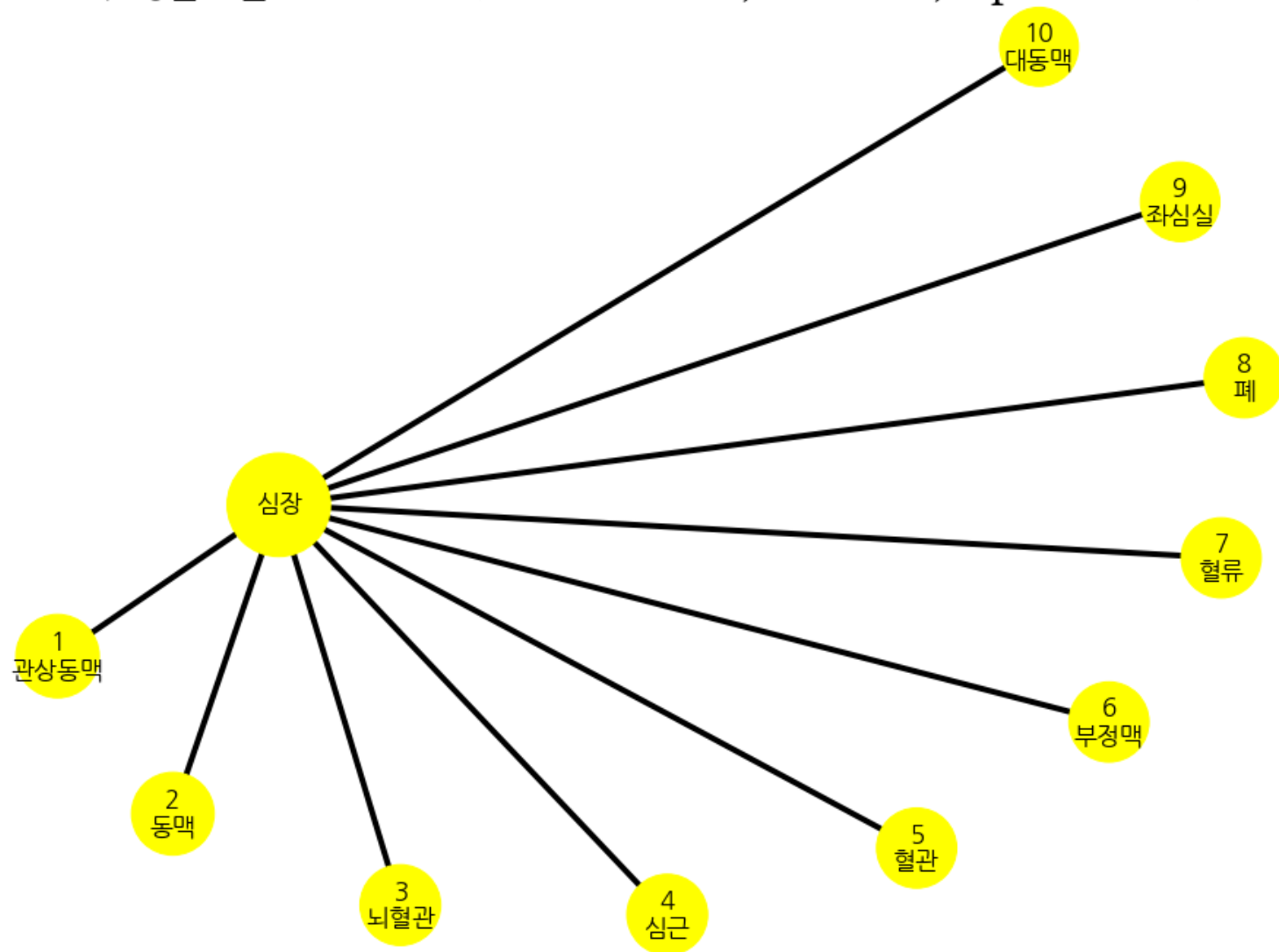
김정은
동지 영도자

"남조선 비하 클러스터"

괴뢰 박근혜
남조선



심장_10_Word2Vec(vocab=19006, size=300, alpha=0.025)



심장_10_FastText(vocab=19006, size=300, alpha=0.025)

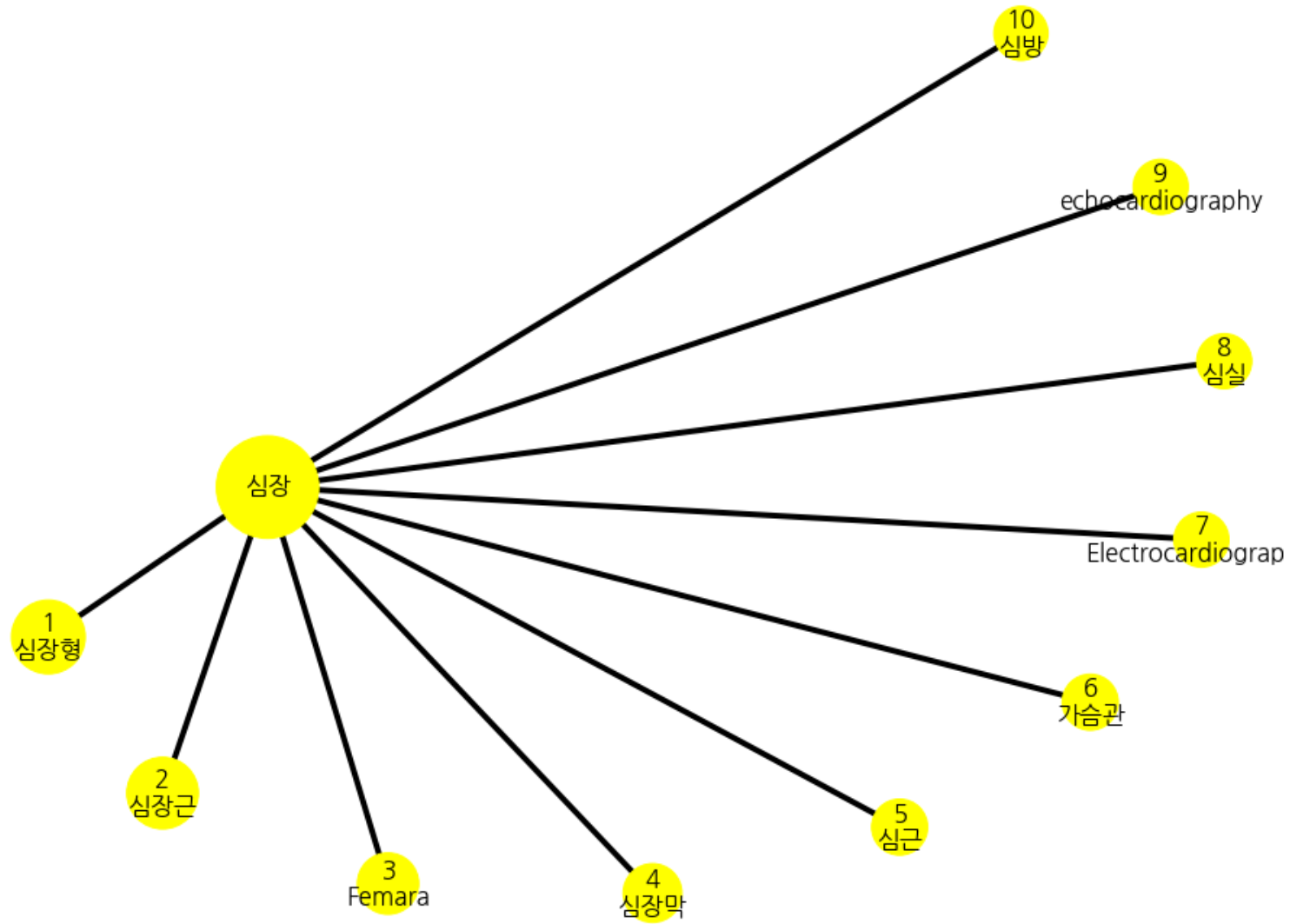
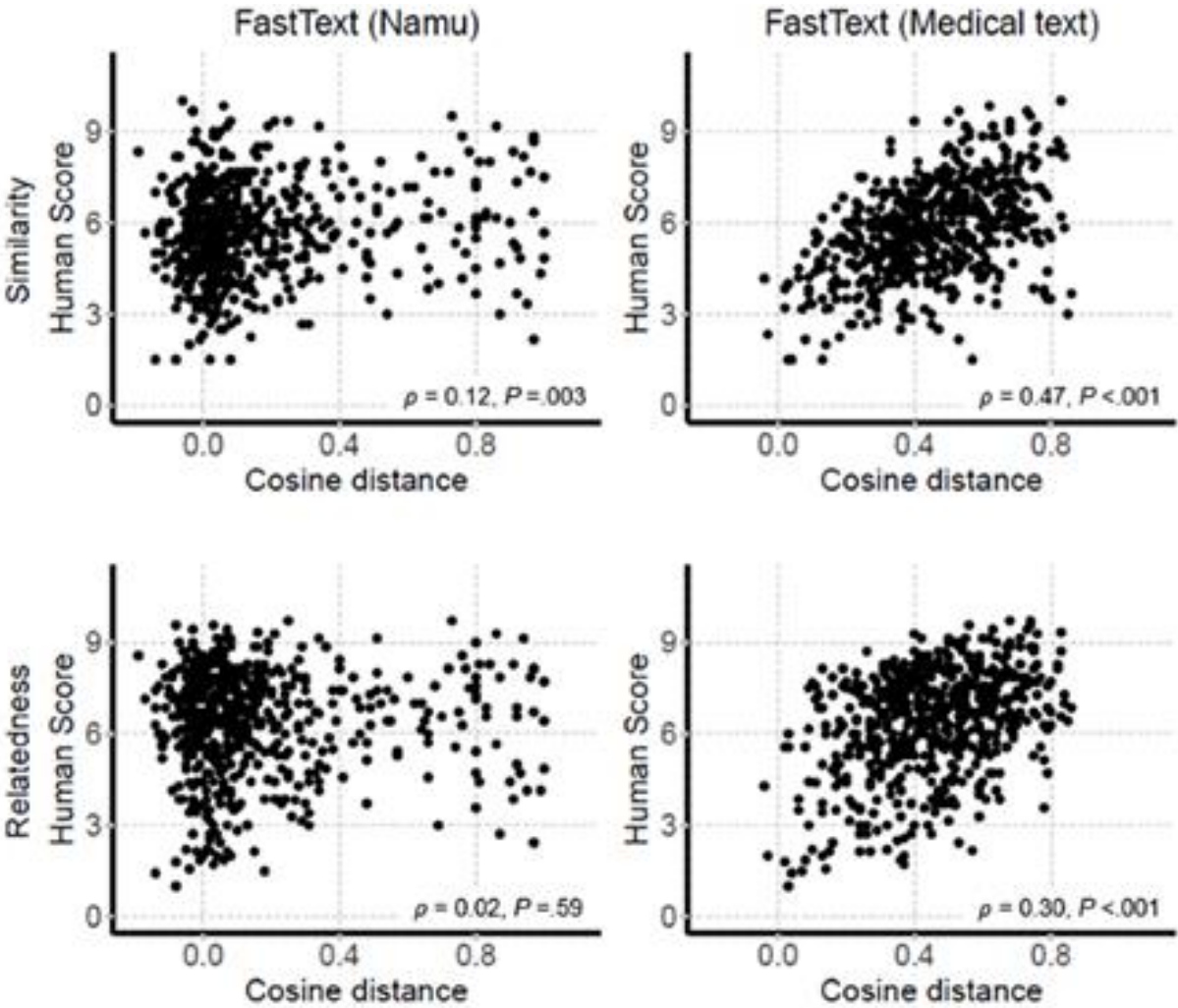


Table 1. Examples of the practice session for human validation in which word pairs of the practice session included medical as well as general terms. Term 1 and Term 2 were presented to the participants. However, their anticipated similarity and relatedness categories were kept hidden.

Term 1	Term 2
책방 (bookstore)	서점 (bookshop)
학교 (school)	경찰서 (police station)
까치 (magpie)	중국어 (the Chinese language)
친구 (friend)	사람 (human)
겨울 (winter)	난로 (heater)
핸드폰 (cell phone)	미술 (art)
심혈관질환 (cardiovascular disease)	관상동맥질환 (coronary artery disease)
암성통증 (cancer pain)	월경통 (menstrual pain)
좌골신경통 (sciatica)	간성혼수 (hepatic coma)
심근경색 (myocardial infarction)	흉통 (chest pain)
세티리진 (cetirizine)	구강건조 (dry mouth)
백반증 (vitiligo)	라미 (lamisil)



비유창성 데이터 임베딩

- 성인 유창성장애 데이터를 워드임베딩에 적용하여 분석
- 연구문제
 - ① 비유창성의 유형(type)과 관련하여, 비유창성 항목들 사이의 분포적 유사성이 존재하는가?
 - ② 비유창성의 유형(type)과 관련하여, 정상적 비유창성과 비정상적 비유창성 항목들 사이의 분포적 차이가 존재하는가?
 - ③ 비유창성의 위치(loci)와 관련하여, 음소 환경에 따른 비유창성의 분포 유형이 존재하는가?

유형	기호	정의
ND	주저 (H)	발화 중간이나 발화 간에 나타나는 1초~3초 정도의 침묵으로 별다른 질적 양상이 동반되지 않은 경우를 말한다.
	간투사 (I)	의미전달내용과 관계없는 낱말이나 구를 말하는 것으로 별다른 질적 양상이 동반되지 않은 경우를 말한다. 예: 그러니까, 글썄, 음, 예, 아, 어 등
	미완성 또는/그리고 수정 (Ur)	발화나 낱말을 끝맺지 않은 경우, 그리고/또는 이미 산출한 말의 발음, 낱말, 통사구조 등을 바꾸어 다시 말하는 경우를 말한다. 미완성/수정(Urv) 유형은 목표음절수에 포함한다. 예: 저-시험에서 저서, 두 고슴도치 등
	반복1 (R1)	다음절낱말이나 구, 어절 등을 1~2회 반복하되, 별다른 질적 양상이 동반되지 않은 경우를 말한다. 예: 승리의승리의, 친구가친구가 등

AD	주저-비정상적 (Ha)	주저함이 3초 이상 지속되거나 시각적 긴장과 같은 질적양상이 동반되는 경우를 말한다.
	간투사-비정상적 (Ia)	간투사를 3회 이상 반복하거나 간투사를 말할 때 시각적 긴장과 같은 질적 양상이 동반되는 경우를 말한다.
	미완성 또는/그리고 수정-비정상적 (URa)	미완성 그리고/또는 수정이 연속적으로 일어나거나, 시각적 긴장과 같은 질적 양상이 동반되는 경우를 말한다.
	반복1-비정상적 (R1a)	다음절날알, 구, 어절 등이 3회 이상 반복되거나 긴장 등을 동반하여 나타나는 경우를 말한다.
	반복2 (R2)	날알보다 작은 단위에서 일어나는 모든 반복을 포함한다. 음소, 음절부분, 음절, 날알부분, 일음절날알의 반복이 이에 속한다. 예: 곰곰이, x자전거, 나날짐승, 기기린 등
	비운율적 발성 (DP)	연장(소리와 공기의 흐름은 유지되나 조음기관의 운동이 멈추는 것), 막힘(공기 또는 목소리의 흐름 및 조음기관의 움직임이 멈추는 것으로 특히 후두의 부적절한 근육 움직임이 동반), 깨진 날알(날알내에서 나타나는 멈춤)을 포함한다. 막힘을 동반한 반복2(R2) 혹은 연장을 동반한 반복2(R2)는 DP와 R2에 각각 표시하여 계산합니다.

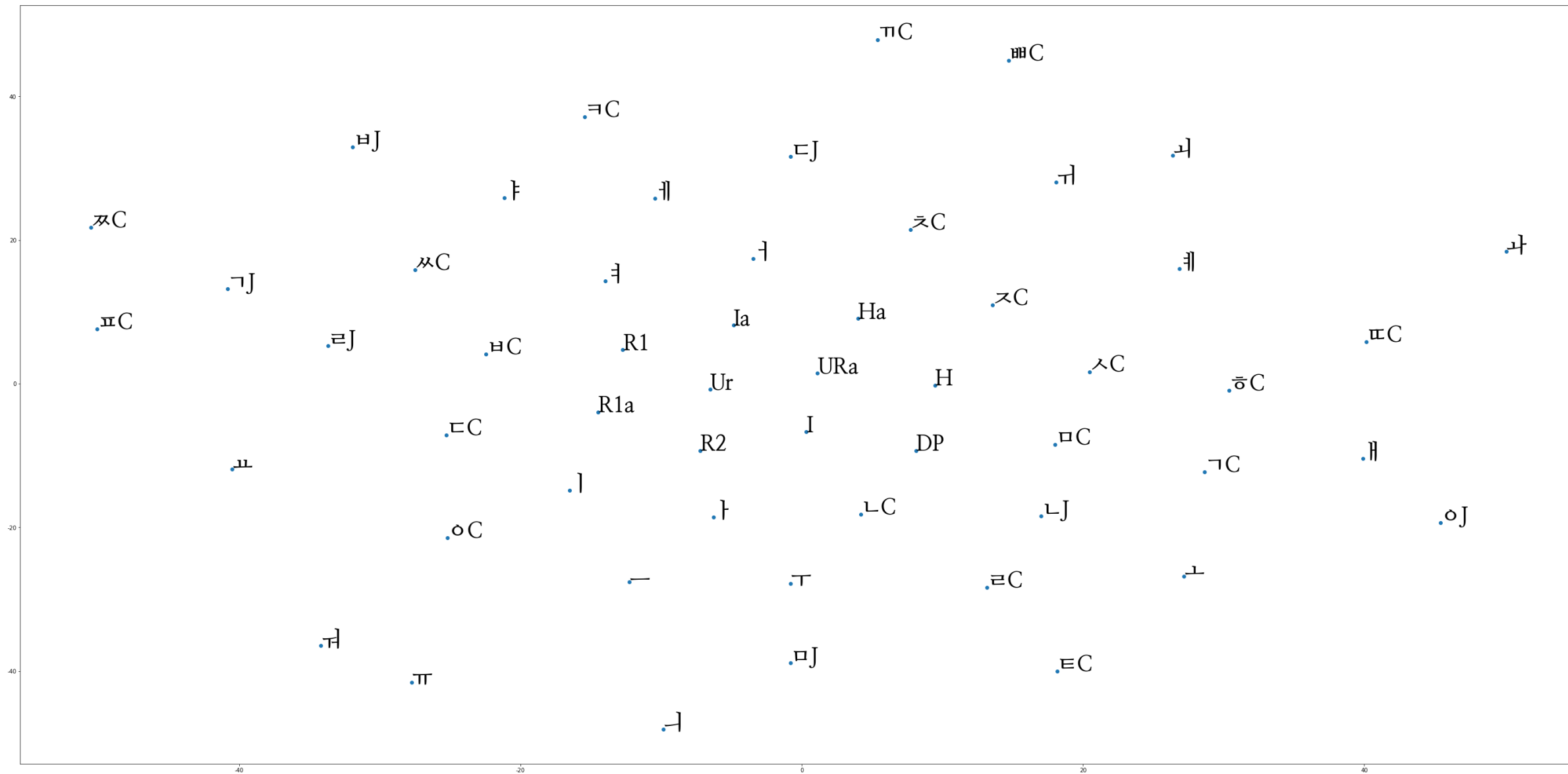
성인 110명 읽기 발화 수집

음절: 112,996개

음소: 236,956개

발화유형 (Fluent=1; SLD = 2; OD=3) SLD+OD도 2번	비유창성 유형 (구체적으로)	발화
--	--------------------	----

2	막힘 1, 음절반복 1	
2	막힘(연장)1, R1 1	2002년 6월 22일 스페DP(연장)인과의 경기는 승부차기까지 R1가는 가는 힘겨운 싸움이였다
1		벽찬 승리의 감동을 이기지 못한 우리는 광화문 거리를 새벽녘에 걸어보았다
3	Ur 1	거리엔 흥분이 가시지 않은 듯 태극기를 흔들거나 Ur당적 경적을 울리며 지나는 차들이 많았다
1		사람들은 붉은 옷을 입고 태극기를 몸에 두르거나 손에 들고 무리지어 대한민국을 외치며 걷고 있었
1		우리는 큰 길에서 벗어나 작은 골목길을 따라 걸었다
1		사람들은 멀리서도 붉은 옷을 입은 사람들만 보면 대한민국을 외치며 두 팔을 번쩍 치켜들었다
1		달려와서 어깨동무하고 펄쩍펄쩍 뛰며 기쁨을 함께 나누는 사람들도 있었다
1		천천히 길을 걸으며 난생 처음 보는 사람들과 기쁨을 나누는 동안 우리는 둥근 공이 가져다 준 일처
1		우리는 세계가 지켜보는 가운데 하나가 되었다
1		김남일 김태영 박지성 설기현 송종국 홍명보 황선홍
1		자랑스런 우리 태극전사들의 이름을 하나씩 마음 속으로 불러 보았다
3	Ur 1	그들의 이름은 이미 Ur개인이 개인의 것이 아니었다
2	R1 1, R2 1	멋지게 꿈을 R1이루어 R2이이루어낸 우리 모두의 것이였다
3	Ur 1	붉은 악마의 소원처럼 Ur뿐 꿈은 이루어진다.



각 유형별 코사인 유사도 - I

H		Ha		I		Ia		Ur		URa		R1		R1a	
Type	sim	Type	sim	Type	sim	Type	sim	Type	sim	Type	sim	Type	sim	Type	sim
R1	0.9655	URa	0.9800	R1a	0.9603	R1a	0.9807	URa	0.9314	Ha	0.9800	R1a	0.9657	Ia	0.9807
R1a	0.9544	Ia	0.9751	Ia	0.9562	URa	0.9779	Ha	0.9260	Ia	0.9779	H	0.9655	Ha	0.9749
Ia	0.9525	R1a	0.9749	Ha	0.9548	Ha	0.9751	Ia	0.9259	R1a	0.9660	Ia	0.9463	URa	0.9660
Ha	0.9237	I	0.9548	URa	0.9501	I	0.9562	H	0.9056	I	0.9501	Ha	0.9238	R1	0.9657
Ur	0.9056	Ur	0.9260	R1	0.9210	H	0.9525	R1a	0.8957	Ur	0.9314	I	0.9210	I	0.9603
URa	0.9039	R1	0.9238	Ur	0.8851	R1	0.9463	R2	0.8944	H	0.9039	R2	0.9185	H	0.9544
R2	0.8858	H	0.9237	R2	0.8788	Ur	0.9259	I	0.8851	R1	0.9018	URa	0.9018	R2	0.9115
I	0.8741	R2	0.9138	H	0.8741	R2	0.8745	R1	0.8815	R2	0.8847	Ur	0.8815	Ur	0.8957
DP	0.6937	DP	0.6592	DP	0.6800	DP	0.6440	□	0.6943	‡	0.6138	DP	0.7103	DP	0.6797
□	0.6586	□	0.5738	□	0.5236	□	0.5812	DP	0.6495	DP	0.5884	⌊	0.6204	‡	0.5638

각 유형별 코사인 유사도 - II

R2		DP	
Type	sim	Type	sim
R1	0.9185	R2	0.7794
Ha	0.9138	R1	0.7103
R1a	0.9115	H	0.6937
Ur	0.8944	I	0.6800
H	0.8858	R1a	0.6797
URa	0.8847	Ha	0.6592
I	0.8788	└	0.6526
la	0.8745	Ur	0.6495
DP	0.7794	la	0.6440
└	0.6650	H	0.6348

무엇부터 배워야 하나요?

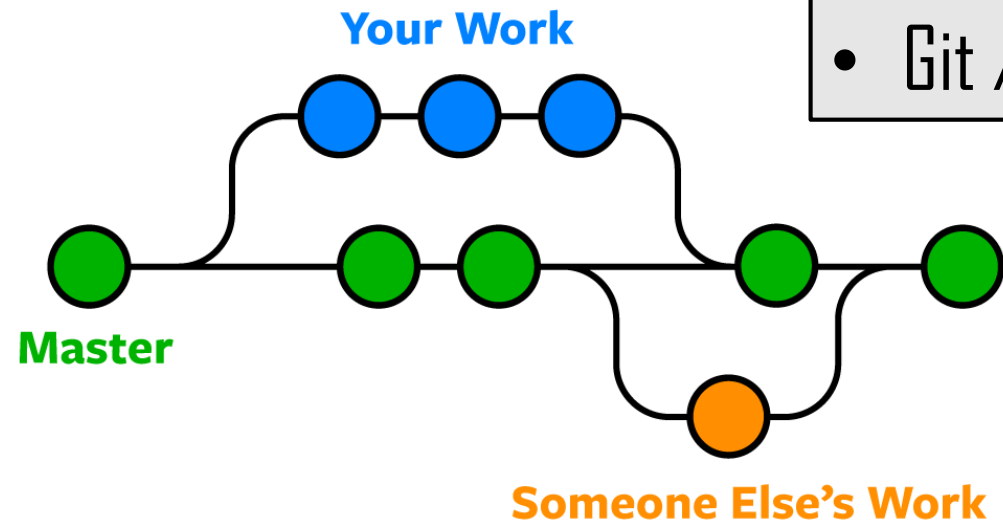
C10	A	B	C	D	E
1					
2		나나 문구점의 일별 판매실적			
3					
4		날짜	연필	지우개	노트
5		07월 01일	50,000	1,200	47,000
6		07월 02일	20,000	5,400	38,000
7		07월 03일	15,500	3,300	20,000
8		07월 04일	7,500	8,900	15,500
9		07월 05일	55,000	5,400	7,500
10		합계	148,000	24,200	128,000
11		평균	29,600	4,840	25,600

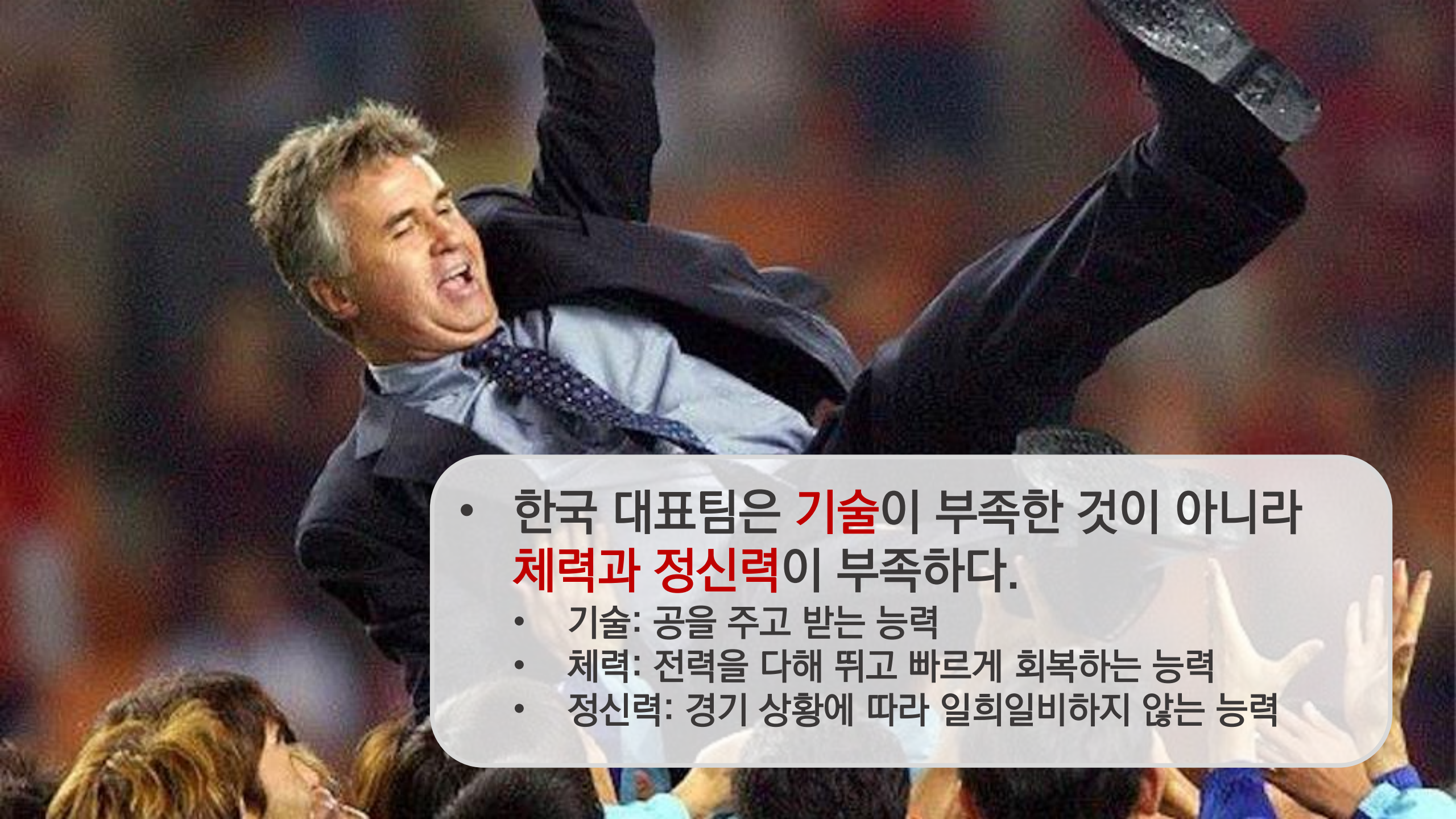
```
File Edit Options Buffers Tools Operate Mark Regexp Immediate Subdir Help
57 (global-set-key (kbd "C-c a") 'screenwriter-action-block)
58 (global-set-key (kbd "C-c d") 'screenwriter-dialog-block)
59 (global-set-key (kbd "C-c t") 'screenwriter-transition)
60 (setq auto-mode-alist (cons '("\\.scp" . screenplay-mode) auto-mode-alist))
61 (setq auto-mode-alist (cons '("\\.md" . markdown-mode) auto-mode-alist))
62
63 ;; w3m setup
64 (setq browse-url-browser-function 'w3m-browse-url)
65 (autoload 'w3m-browse-url "w3m" "Ask a WWW browser to show a URL." t)
66 (global-set-key "C-xm" 'browse-url-at-point)
67 (setq w3m-use-cookies t)
68
69 ;; auto-complete
70 ;; install by running emacs and doing an m-x load-file.el
71 ;; load ~/.emacs.d/auto-complete/etc/install.el
-- .emacs 21% L68 (Emacs-Lisp AC Abbrev)
8 ** <2021-09-18 1300-1600>
9 * Grocery
10 :CATEGORY: Food
11 ** TODO Artichokes
12 ** TODO Bagels
13 - Flour
14 - Baking soda
15 - Rock salt
16 ** Pretzels
17
18
--*- List.org Bot L12 (Org U:%%- a-compat32 2% L5
```

- Excel
- Text Editor

```
dnyce@LinuxShellTips: ~
dnyce@LinuxShellTips: ~
dnyce@LinuxShellTips: ~
dnyce@LinuxShellTips:~$ script --t=timingfile.txt -a typescript.txt
Script started, output log file is 'typescript.txt', timing file is 'timingfile.txt'.
dnyce@LinuxShellTips:~$ ls
Desktop  Documents  LinuxShellTips  Pictures  RandNumGen.sh  Templates  typescript.txt
dir1     Downloads  Music           Public   snap           timingfile.txt Videos
dnyce@LinuxShellTips:~$ cat RandNumGen.sh
#!/bin/bash
#LinuxShellTips Script to Generate a Random Number
RANDOM=$$
for i in RANDOM
do
    echo $RANDOM
done
dnyce@LinuxShellTips:~$ rm dir1
rm: cannot remove 'dir1': Is a directory
dnyce@LinuxShellTips:~$ rm -r dir1
dnyce@LinuxShellTips:~$
exit
Script done.
dnyce@LinuxShellTips:~$
```

- Linux Terminal
- Git / Git Hub



- 
- A man in a dark suit and light blue shirt is being lifted high into the air by a crowd of people. He has a joyful expression, with his mouth open as if shouting or cheering. The background is a blurred crowd of people, some with their hands raised, suggesting a celebratory event like a sports victory. The lighting is warm and focused on the man.
- 한국 대표팀은 **기술**이 부족한 것이 아니라 **체력과 정신력**이 부족하다.
 - 기술: 공을 주고 받는 능력
 - 체력: 전력을 다해 뛰고 빠르게 회복하는 능력
 - 정신력: 경기 상황에 따라 일희일비하지 않는 능력

정규표현식:

언론사별 300개씩, 20개 이상의 언론사

대단한 착각:

데이터만 많으면 딥러닝이 잘 작동할 것이다.

불편한 진실:

정제된 데이터만 많으면 딥러닝이 잘 작동할 것이다.

EXCEL 부터 연마하자!

 If문

 vlookup

 부분합

 피벗테이블 보고서

 틀 고정

 필터 사용

 빈칸 안 만들기

 셀 서식 적용

Text Editor



메모장



gedit



EditPlus



Notepad++



Visual Studio / eclipse



Sublime Text



emacs



VS Code

Text Editor – 기능

 정렬

 인코딩 변환

 정규표현식

 찾기/바꾸기

 Syntax Highlighting

 매크로

 SSH

 indent