WHAT I HAVE LEARNED AFTER DOING 50+ KAGGLE COMPETITIONS

DATA + ADA

JOHN PARK

- 1시간 안에 알려 드릴 테크니컬은 별로 없음.

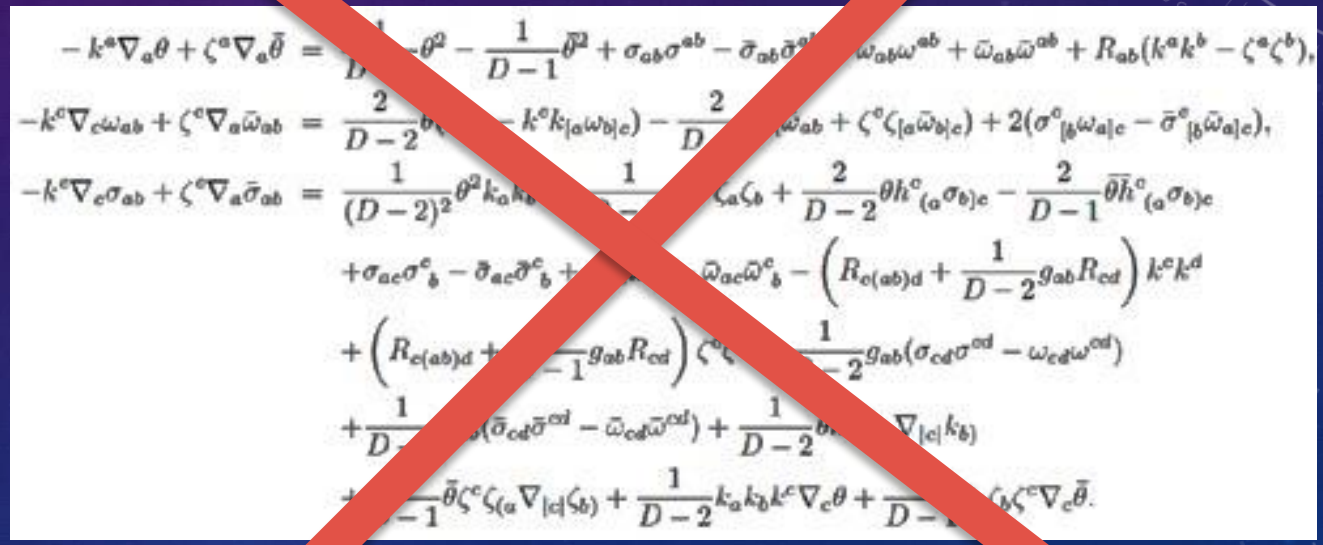
- 그냥 인터넷 보는게 훨씬빠름.
  - 페북/TensorFlow KR
  - 페북/AI Korea.
  - DataTau
  - /r/MachineLearning

전달하고 싶은 내용은..

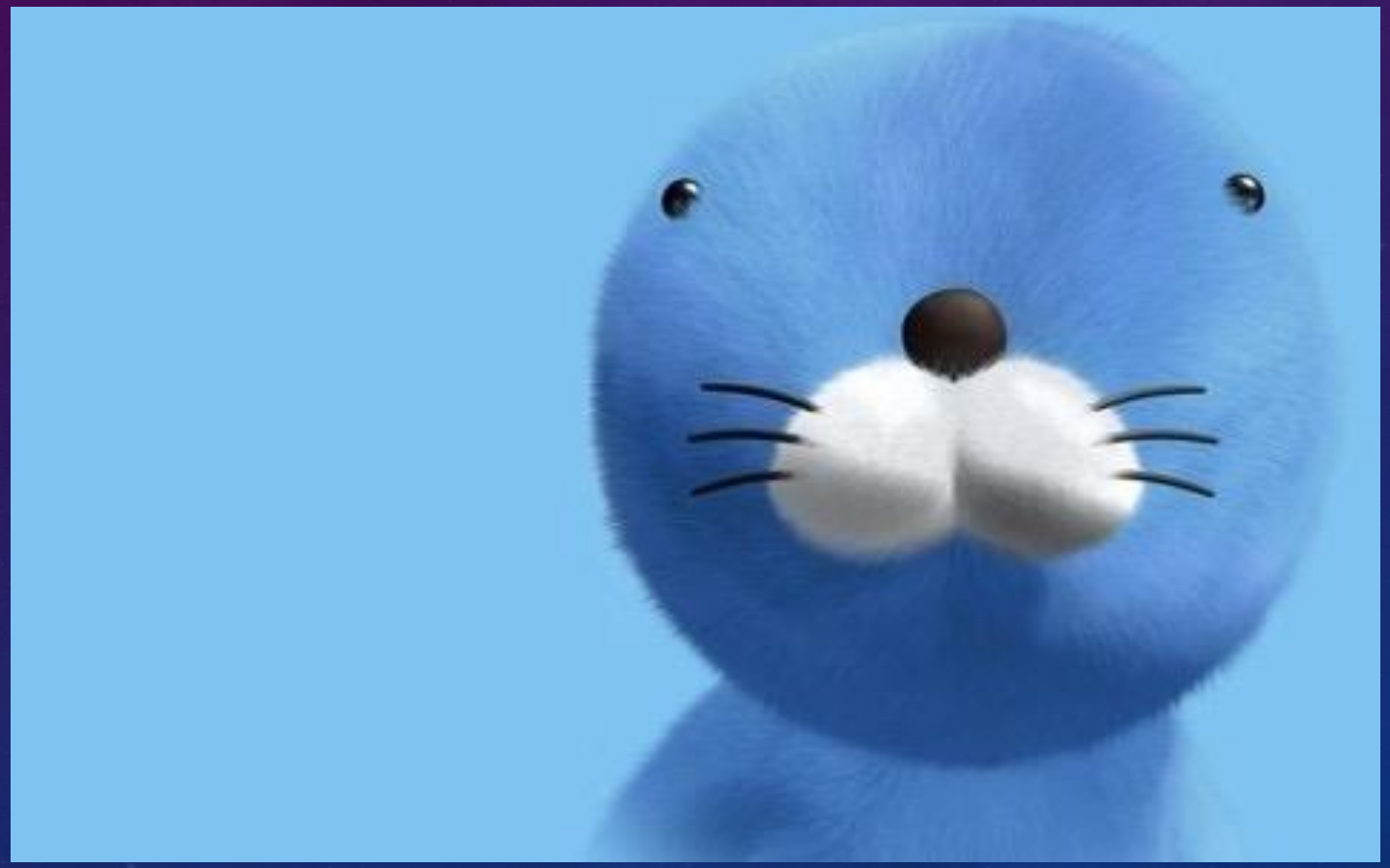
- 머신러닝을 배우려면 어떻게 시작하나요?

가장 중요한건

- NO 수식



수식 대신에, 보노보노 같은 사진들

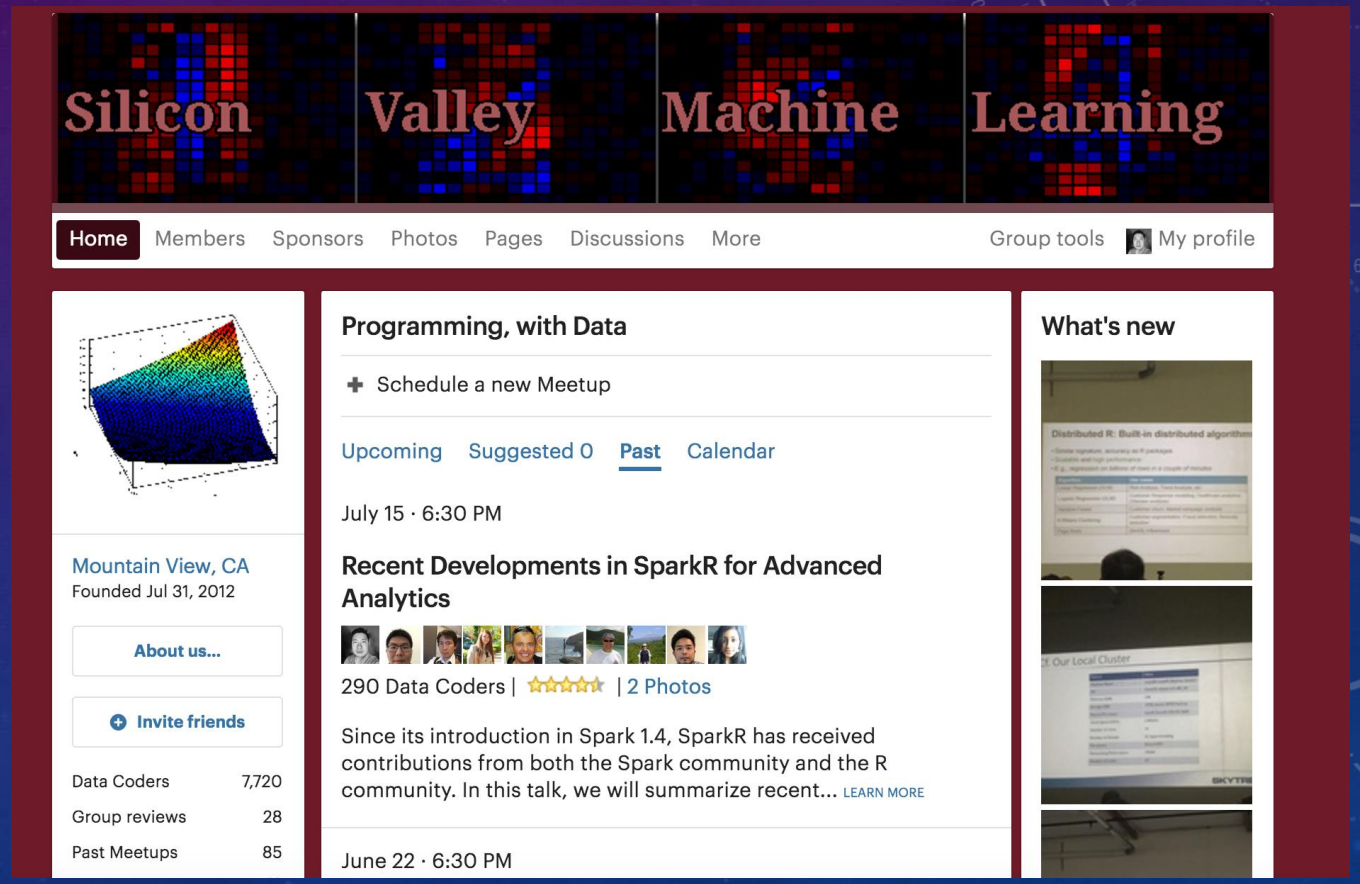



뭣이 중헌디?!
kyokyo's 세모아

- 100+ Slides, 40 Minutes

SILICON VALLEY MACHINE LEARNING MEET-UP

- organizer for the last 3 years
- grew it from 800 to 7800 members
- the largest Machine Learning group in the world.



STARTED WITH FEW GUYS

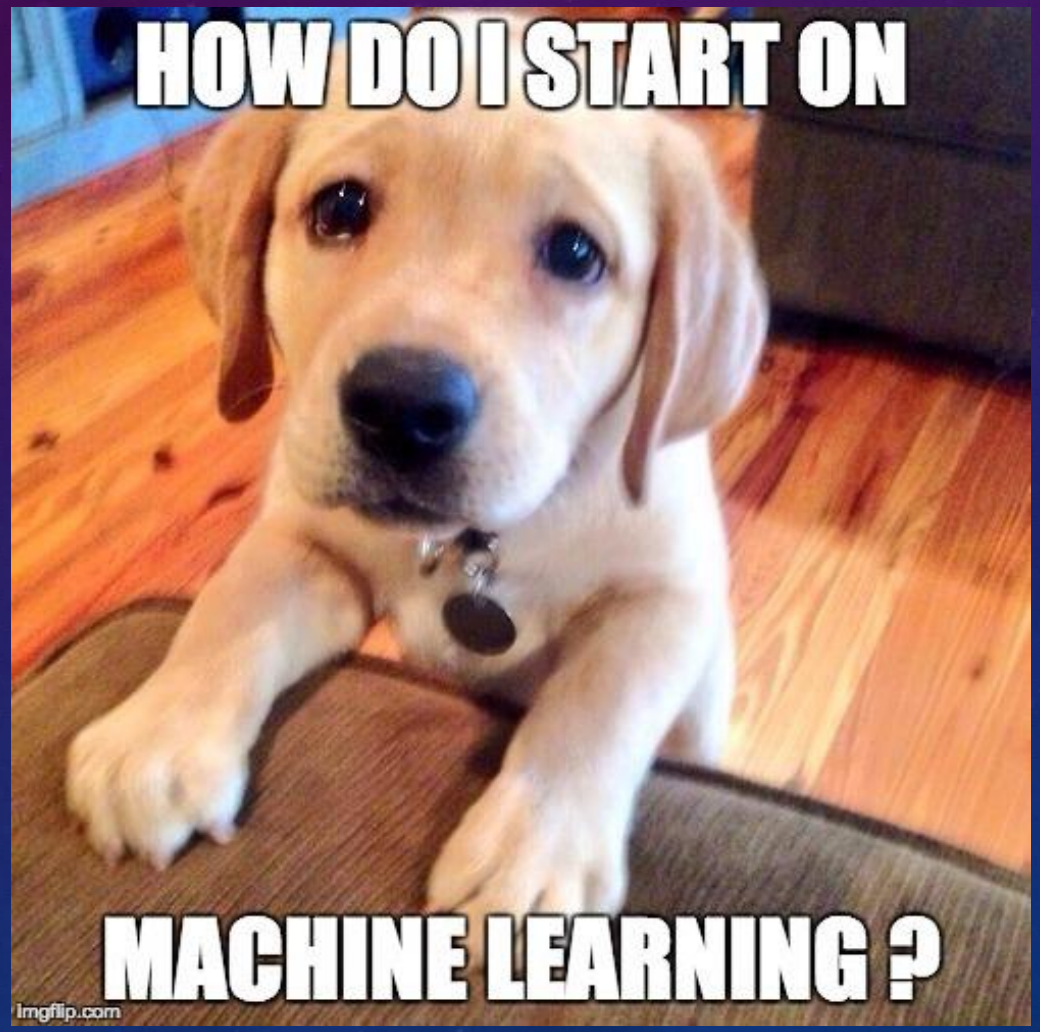
- just few guys, met at HackerDojo (해커스페이스).
- we hijacked the group, without permission.

TO GET TOGETHER AND HACK TOGETHER

START-UP PRESENTATIONS

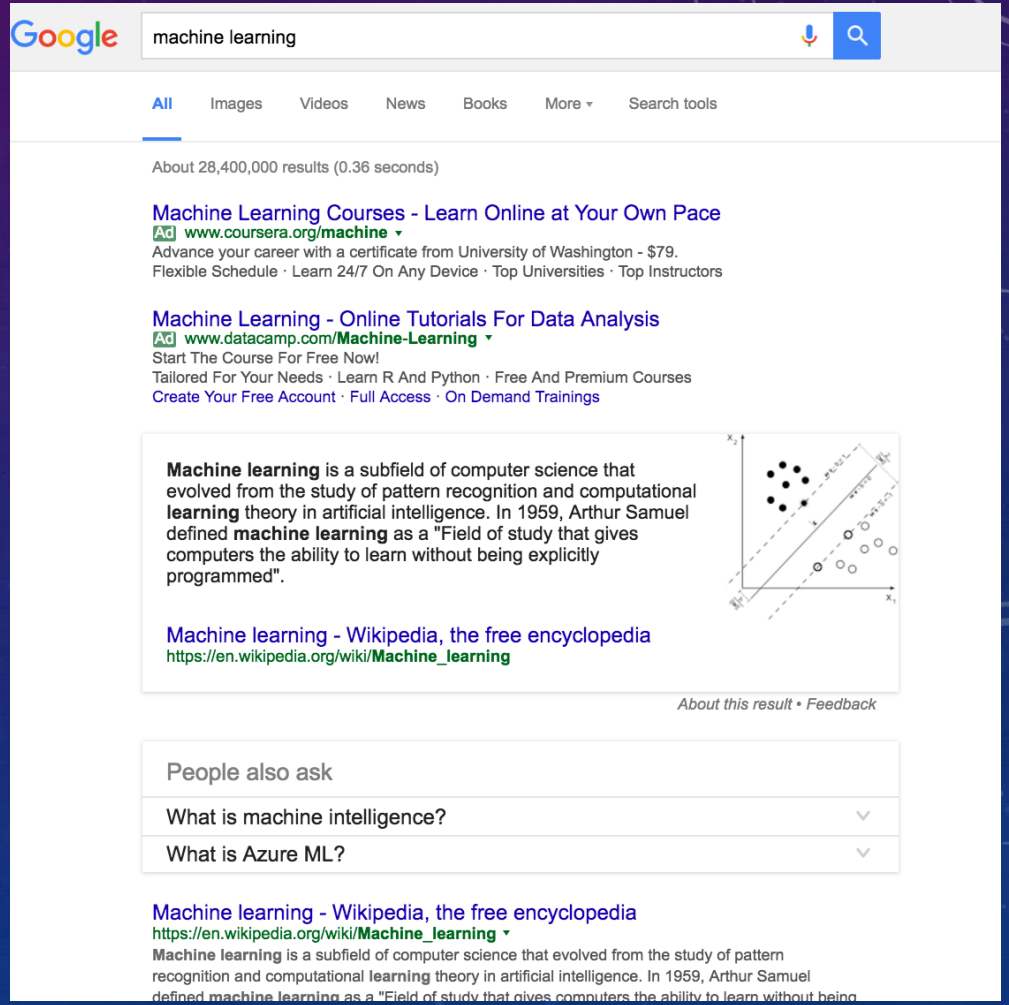
RECRUITING EVENT

ONE UNIVERSAL QUESTION



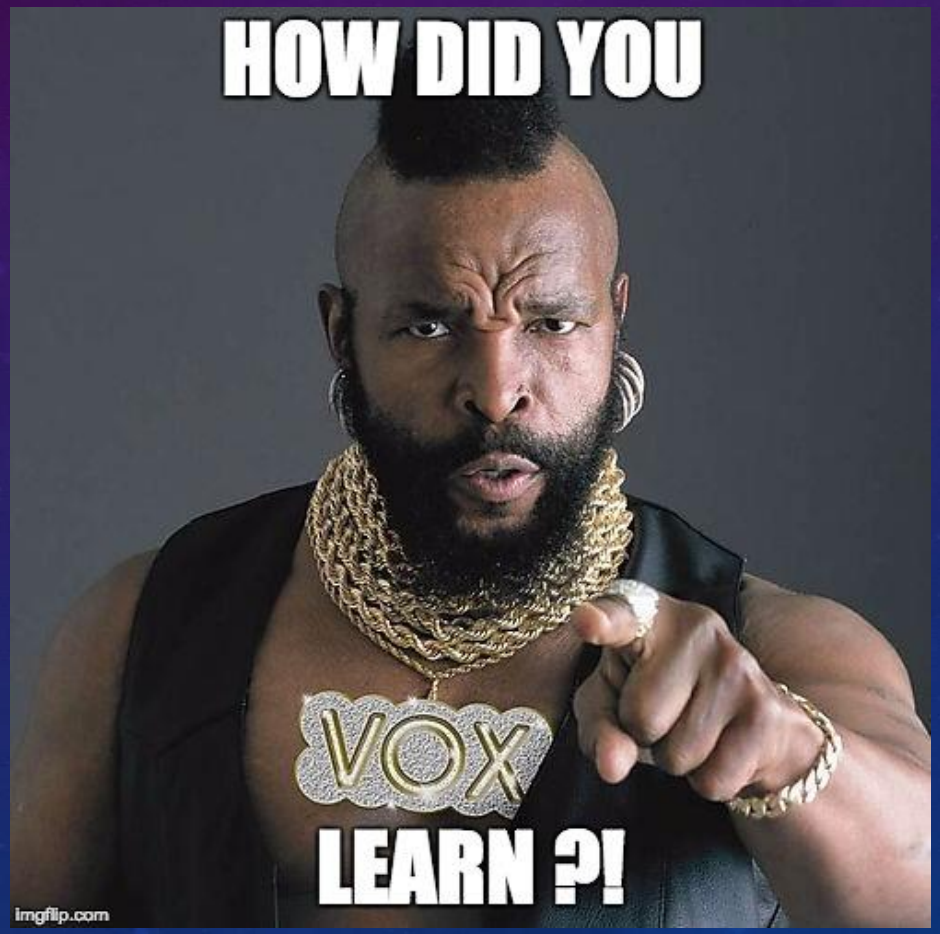
ONE UNIVERSAL ANSWER

- 구글에다 “machine learning” 치고.
- 처음 10 페이지에 있는 링크를 다 읽고...
- … 3 layer 정도만 따라가봐.



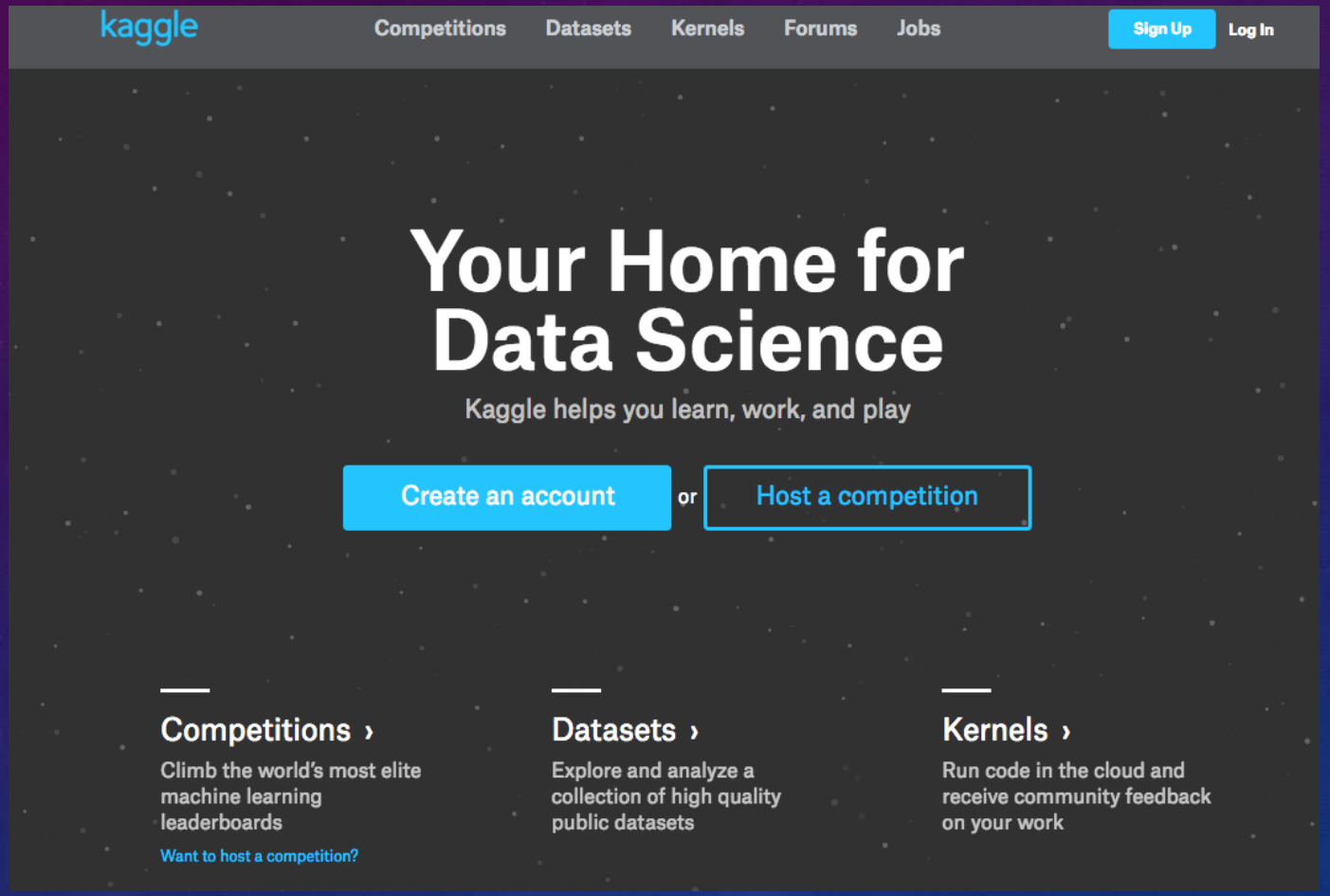

대답이
너무 간단해

ONE UNIVERSAL FOLLOW UP QUESTION



ONE UNIVERSAL FOLLOW UP ANSWER

ONE UNIVERSAL EXCUSE

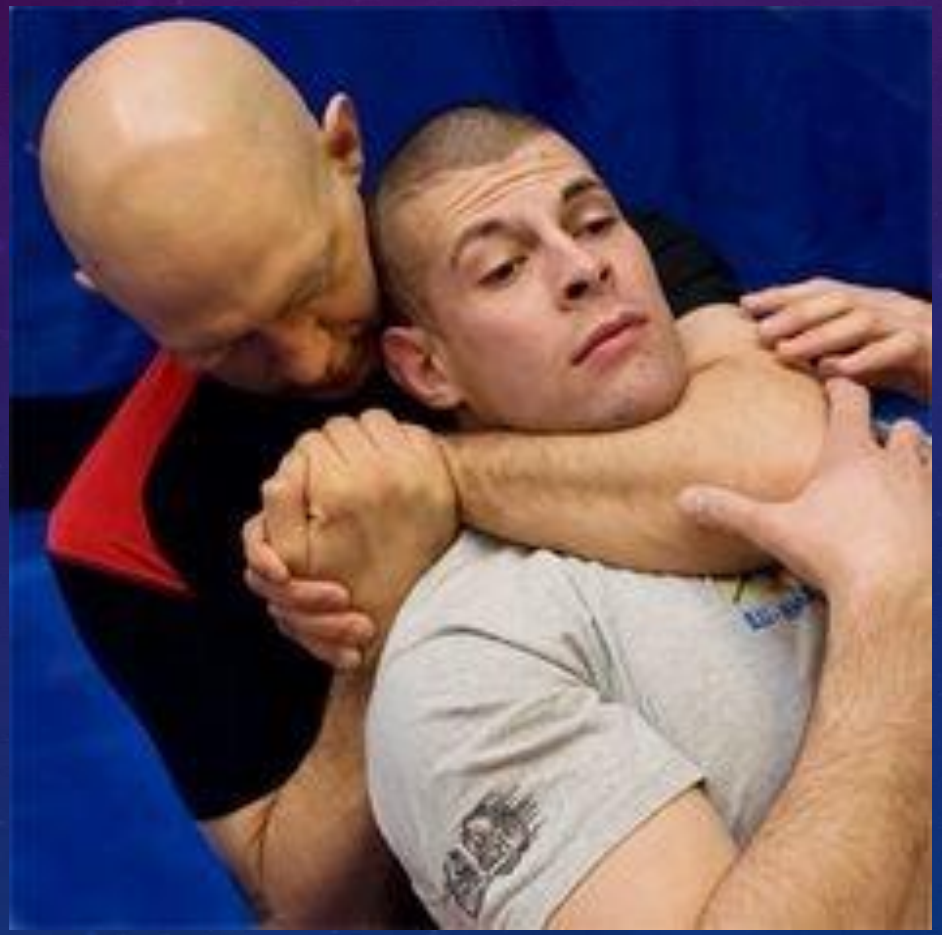

kaggle
Competitions
Datasets
Kernels
Forums
Jobs
Sign Up
Log In
Your Home for Data Science
Kaggle helps you learn, work, and play
Create an account
or
Host a competition
Competitions ›
Climb the world's most elite machine learning leaderboards
Want to host a competition?
Datasets ›
Explore and analyze a collection of high quality public datasets
Kernels ›
Run code in the cloud and receive community feedback on your work

SIMILAR TO IMAGENET, BUT ON DIVERSE TOPICS

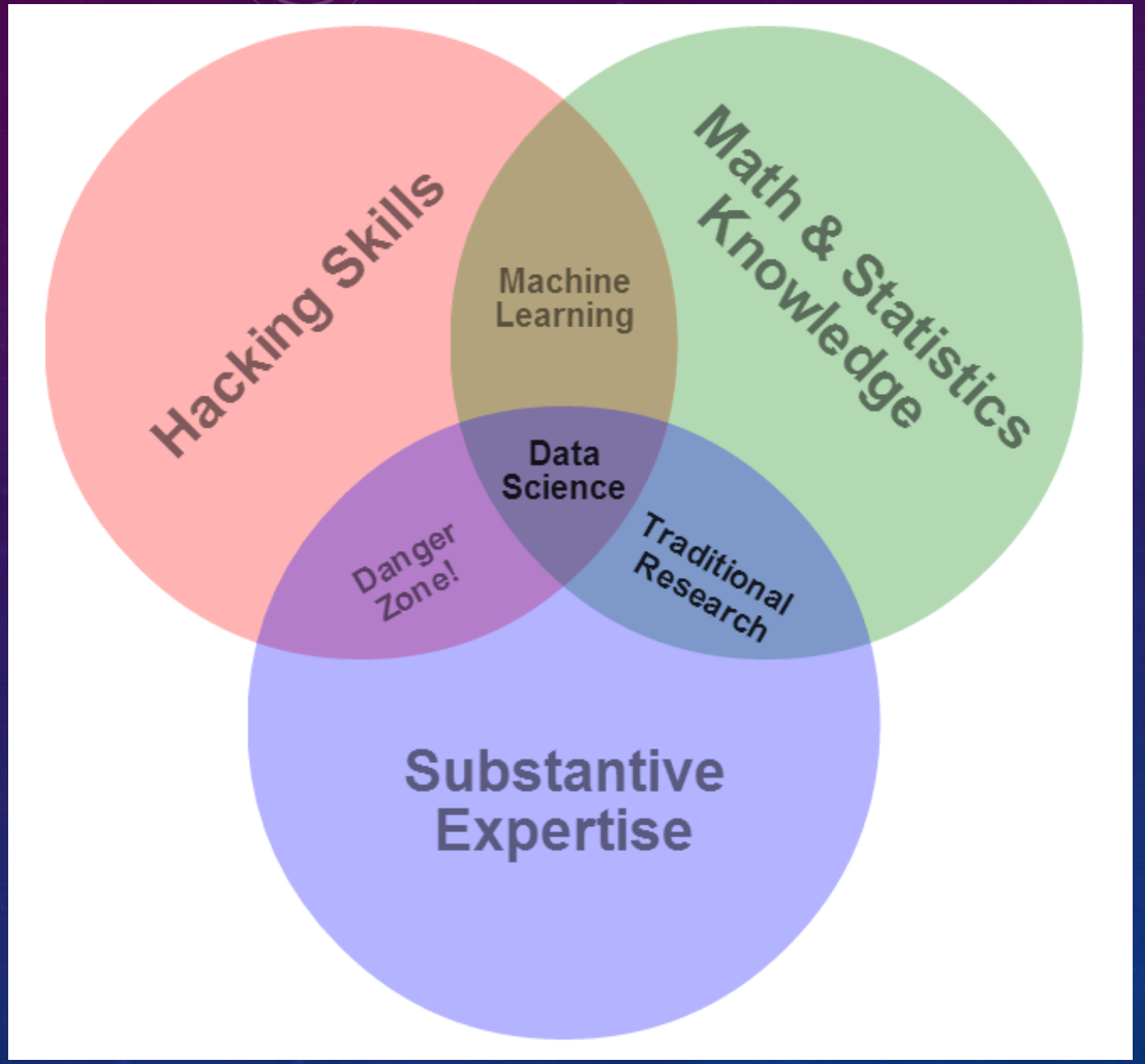
NO RESTRICTION ON METHODS

NOT THESE

THIS WORKS.



- Kaggle is not everything.

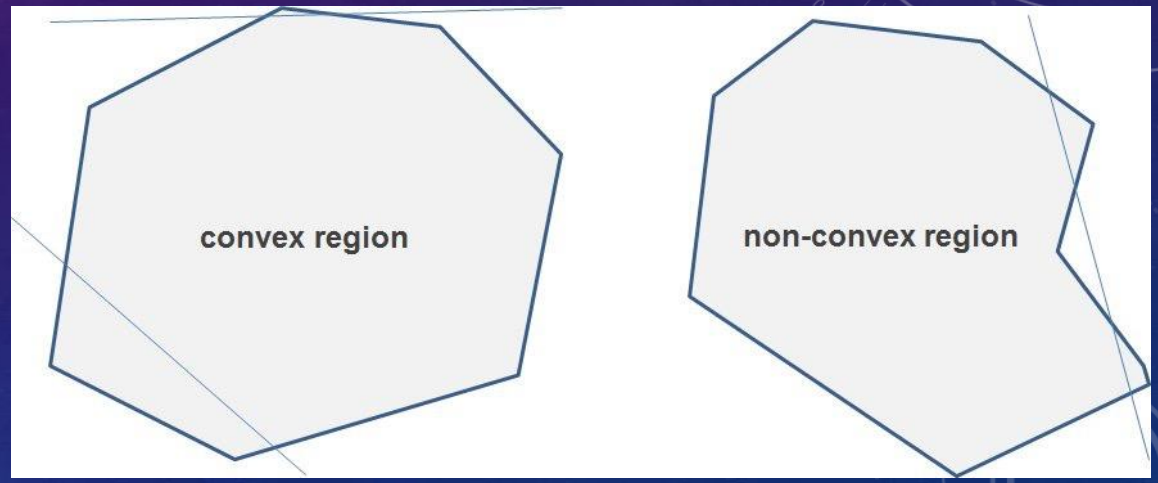

Hacking Skills
Math & Statistics Knowledge
Substantive Expertise
Machine Learning
Data Science
Danger Zone!
Traditional Research

NOT COVERED

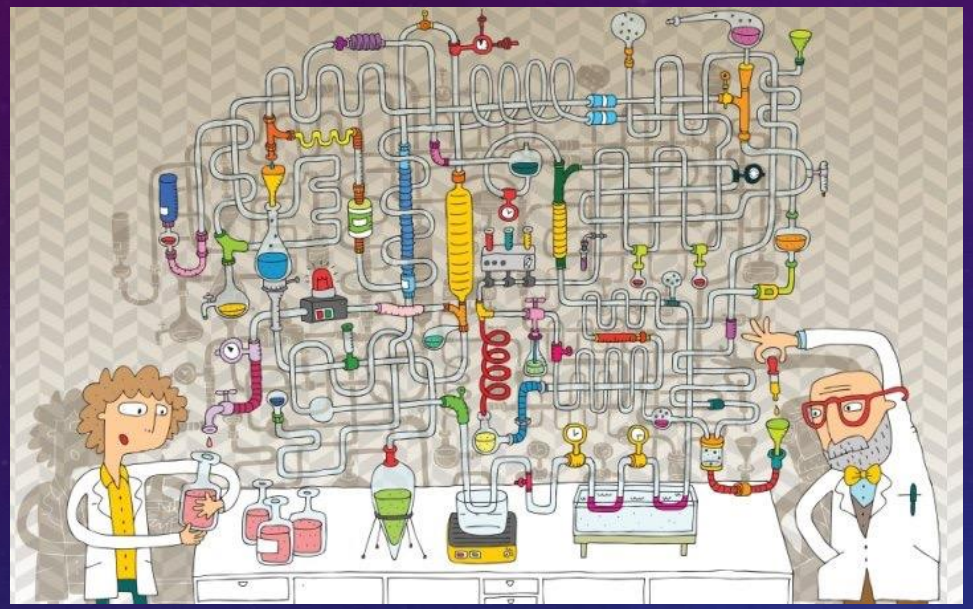
- Project scoping.
- Error metric.
- Data crawling
- Data cleaning.

MORE PRACTICAL THAN ACADEMIA

- to work with real world data. no more toy data.
- any topic you like.
- to make a working model, and compare with others. (error metric yourself!)



WHY I LIKE IT



+

- Work with people who are f*cking brilliant. (이런 인간들이 존재하는지 몰랏다.)

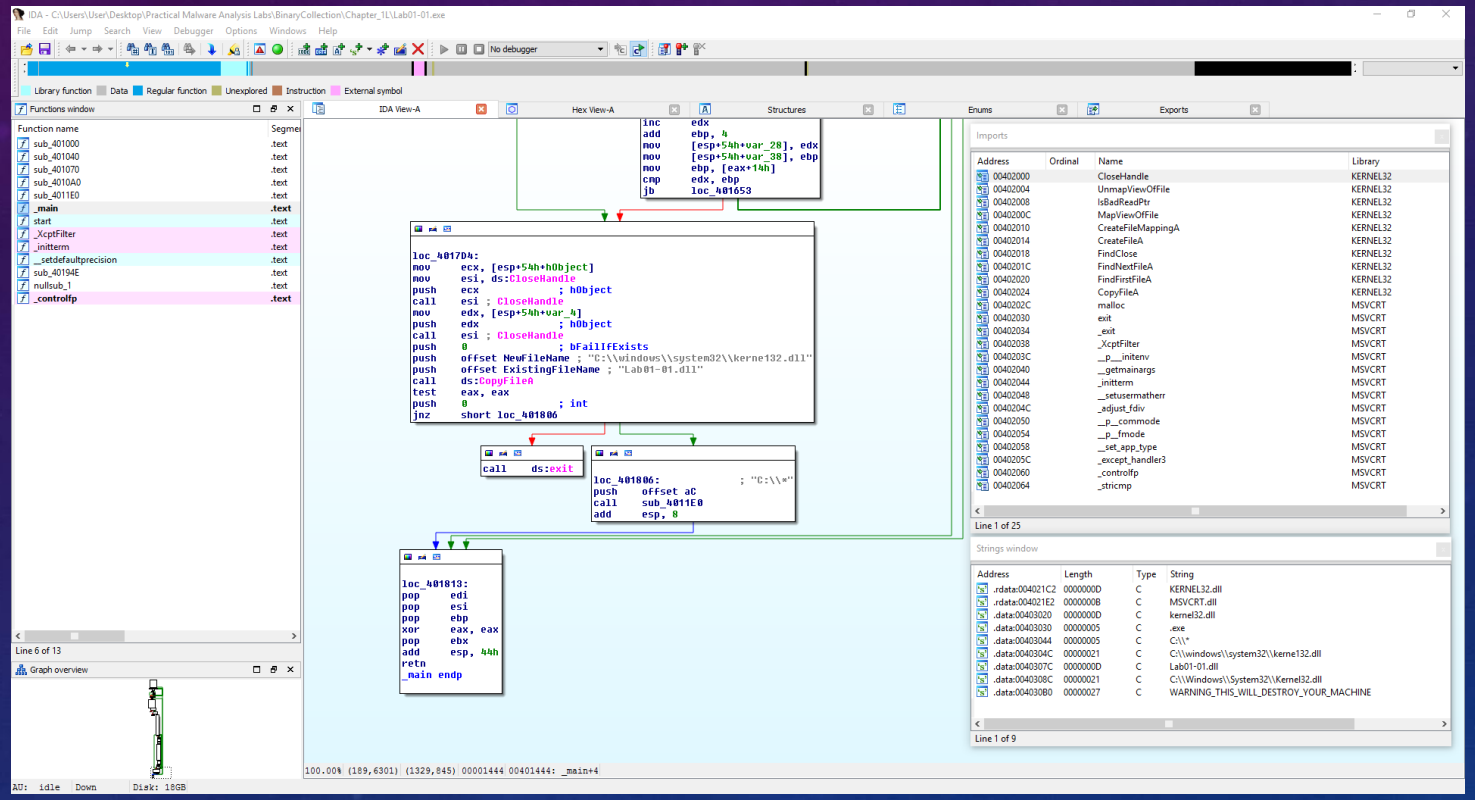
EVERYONE’S LIFE IS DIFFERENT

TL;DR VERSION

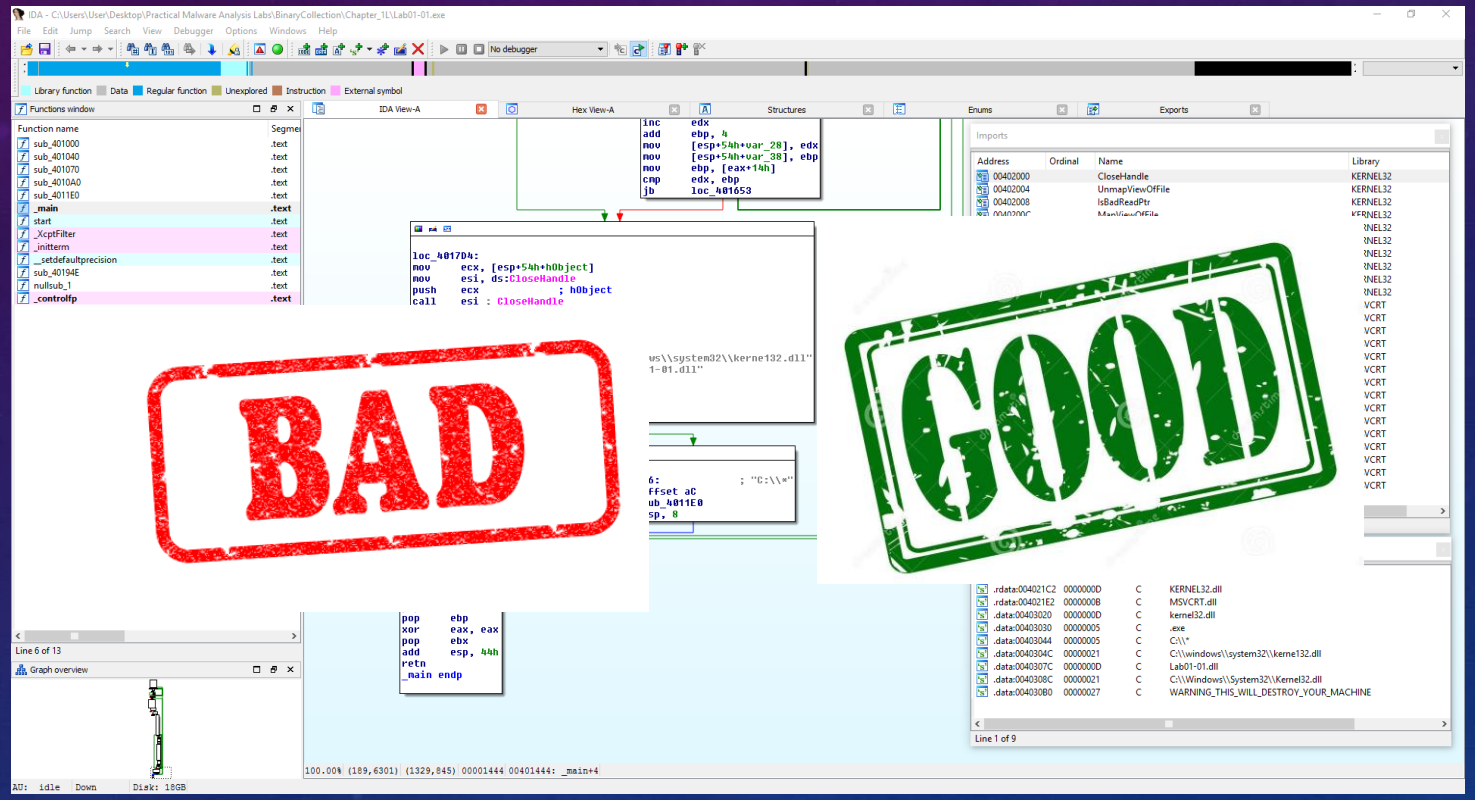
- 1. 놀려고
- 2. 놀고 있어서

보안업계 일함.

악성코드 분석

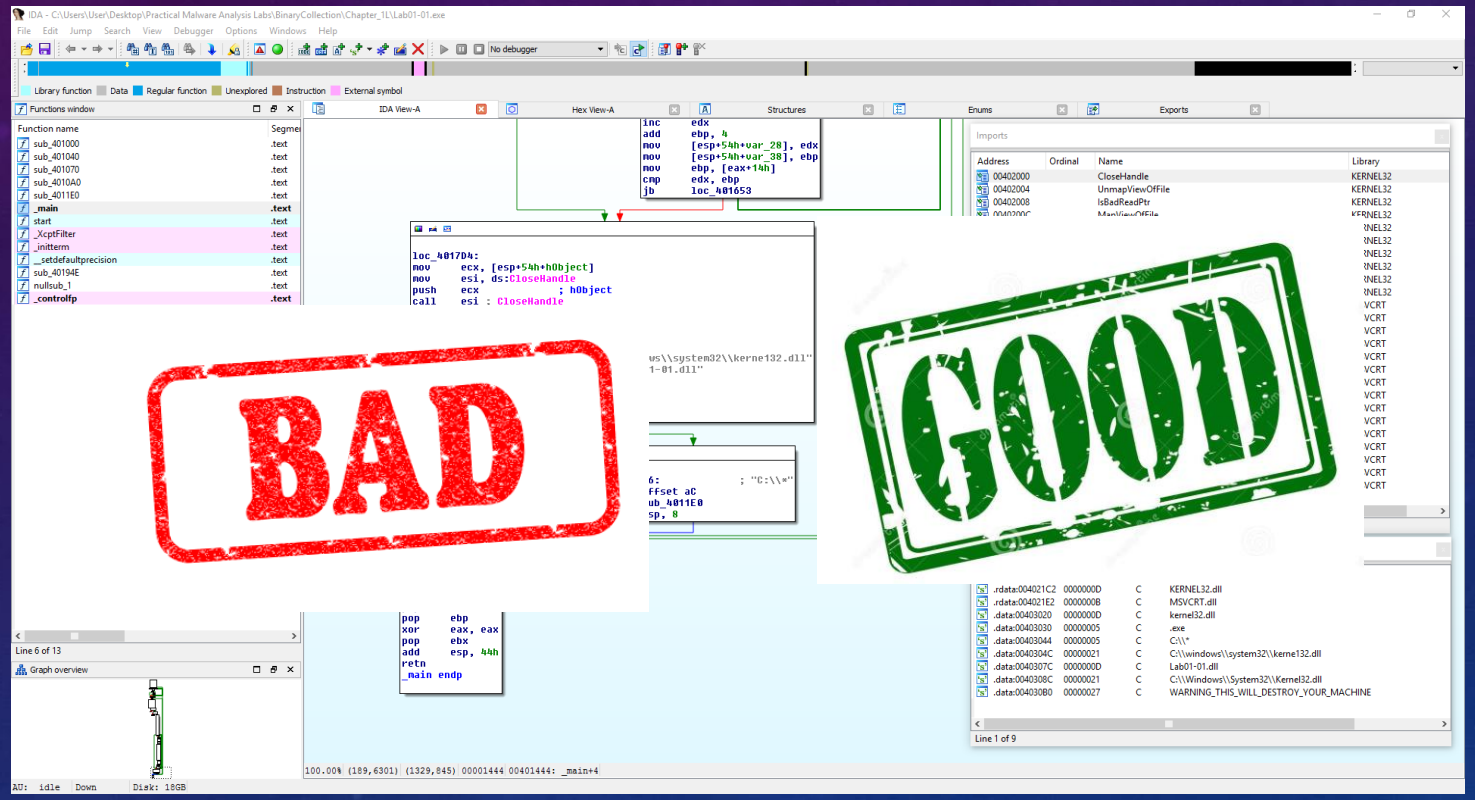

IDA - C:\Users\User\Desktop\Practical Malware Analysis Labs\BinaryCollection\Chapter_1L\Lab01-01.exe
File Edit Jump Search View Debugger Options Windows Help
No debugger
Library function Data Regular function Unexplored Instruction External symbol
Functions window
Function name
sub_401000
sub_401040
sub_401070
sub_4010A0
sub_4011E0
_main
start
_XcptFilter
_initterm
__setdefaultprecision
sub_40194E
nullsub_1
_controlfp
Line 6 of 13
Graph overview
IDA View-A
Hex View-A
Structures
Enums
Exports
inc edx
add ebp, 4
mov [esp+54h+var_28], edx
mov [esp+54h+var_38], ebp
mov ebp, [eax+14h]
cmp edx, ebp
jb loc_401653
loc_4017D4:
mov ecx, [esp+54h+hObject]
mov esi, ds:CloseHandle
push ecx ; hObject
call esi ; CloseHandle
mov edx, [esp+54h+var_4]
push edx ; hObject
call esi ; CloseHandle
push 0 ; bFailIfExists
push offset NewFileName ; "C:\\windows\\system32\\kerne132.dll"
push offset ExistingFileName ; "Lab01-01.dll"
call ds:CopyFileA
test eax, eax
push 0 ; int
jnz short loc_401806
call ds:exit
loc_401806: ; "C:*"
push offset aC
call sub_4011E0
add esp, 8
loc_401813:
pop edi
pop esi
pop ebp
xor eax, eax
pop ebx
add esp, 44h
retn
_main endp
Imports
Address Ordinal Name Library
00402000 CloseHandle KERNEL32
00402004 UnmapViewOfFile KERNEL32
00402008 IsBadReadPtr KERNEL32
0040200C MapViewOfFile KERNEL32
00402010 CreateFileMappingA KERNEL32
00402014 CreateFileA KERNEL32
00402018 FindClose KERNEL32
0040201C FindNextFileA KERNEL32
00402020 FindFirstFileA KERNEL32
00402024 CopyFileA KERNEL32
0040202C malloc MSVCRT
00402030 exit MSVCRT
00402034 _exit MSVCRT
00402038 _XcptFilter MSVCRT
0040203C __p___initenv MSVCRT
00402040 __getmainargs MSVCRT
00402044 _initterm MSVCRT
00402048 __setusermatherr MSVCRT
0040204C _adjust_fdiv MSVCRT
00402050 __p__commode MSVCRT
00402054 __p__fmode MSVCRT
00402058 __set_app_type MSVCRT
0040205C _except_handler3 MSVCRT
00402060 _controlfp MSVCRT
00402064 _stricmp MSVCRT
Line 1 of 25
Strings window
Address Length Type String
.rdata:004021C2 0000000D C KERNEL32.dll
.rdata:004021E2 0000000B C MSVCRT.dll
.data:00403020 0000000D C kernel32.dll
.data:00403030 00000005 C .exe
.data:00403044 00000005 C C:*
.data:0040304C 00000021 C C:\\windows\\system32\\kerne132.dll
.data:0040307C 0000000D C Lab01-01.dll
.data:0040308C 00000021 C C:\\Windows\\System32\\Kernel32.dll
.data:004030B0 00000027 C WARNING_THIS_WILL_DESTROY_YOUR_MACHINE
Line 1 of 9
100.00% (189,6301) (1329,845) 00001444 00401444: _main+4
AU: idle Down Disk: 18GB

악성코드 분석



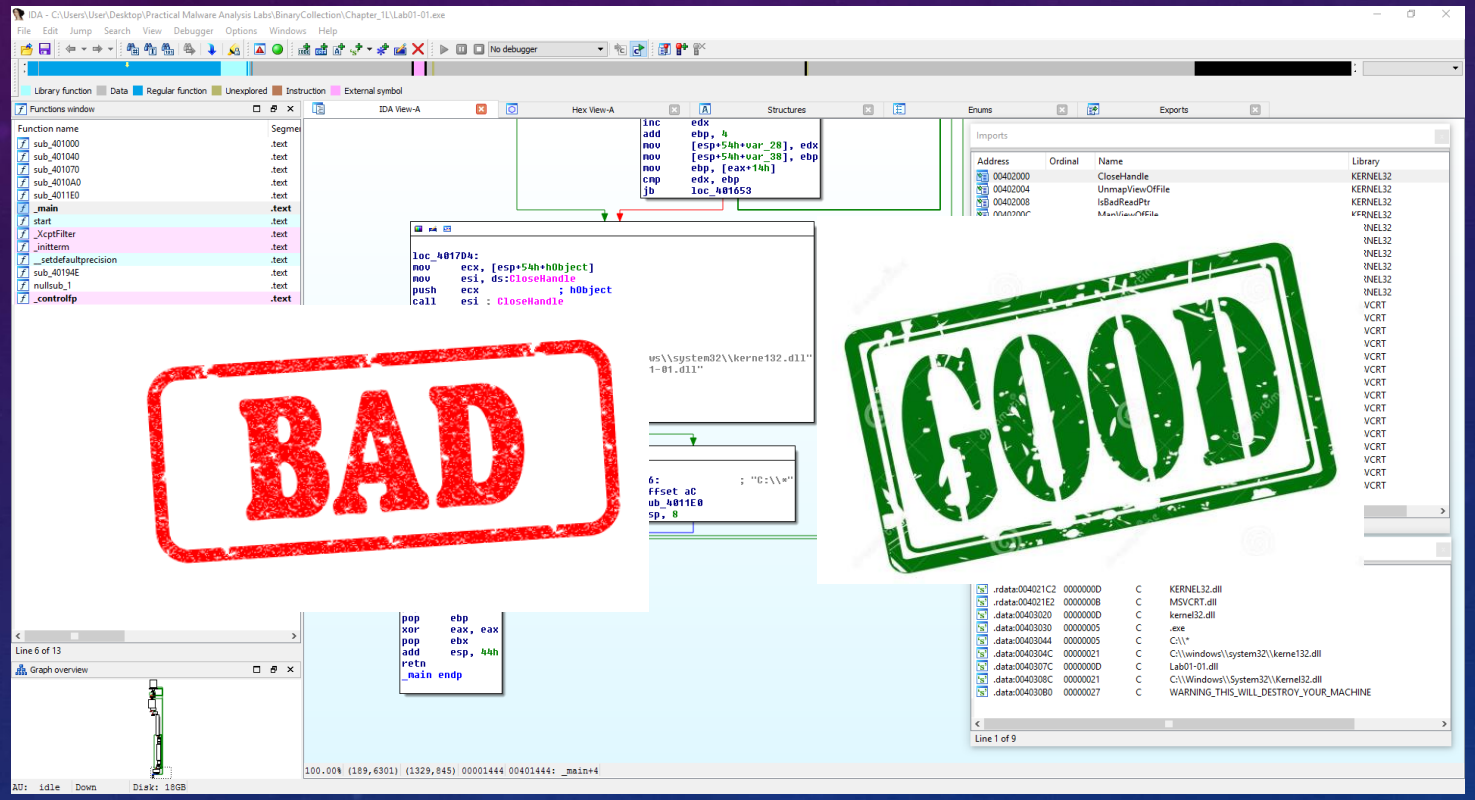

오늘도...
imgflip.com

악성코드 분석




내일도...
imgflip.com

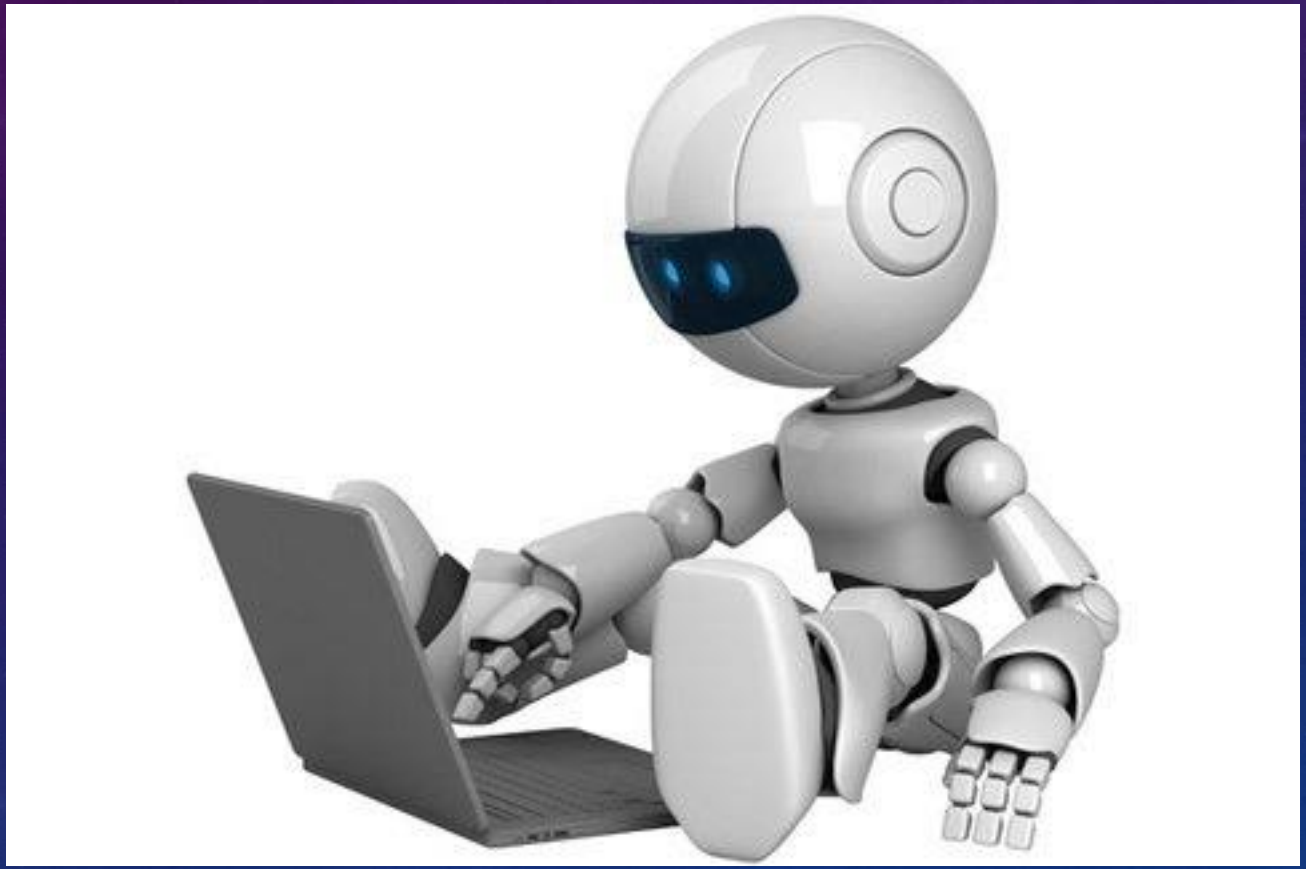
악성코드 분석

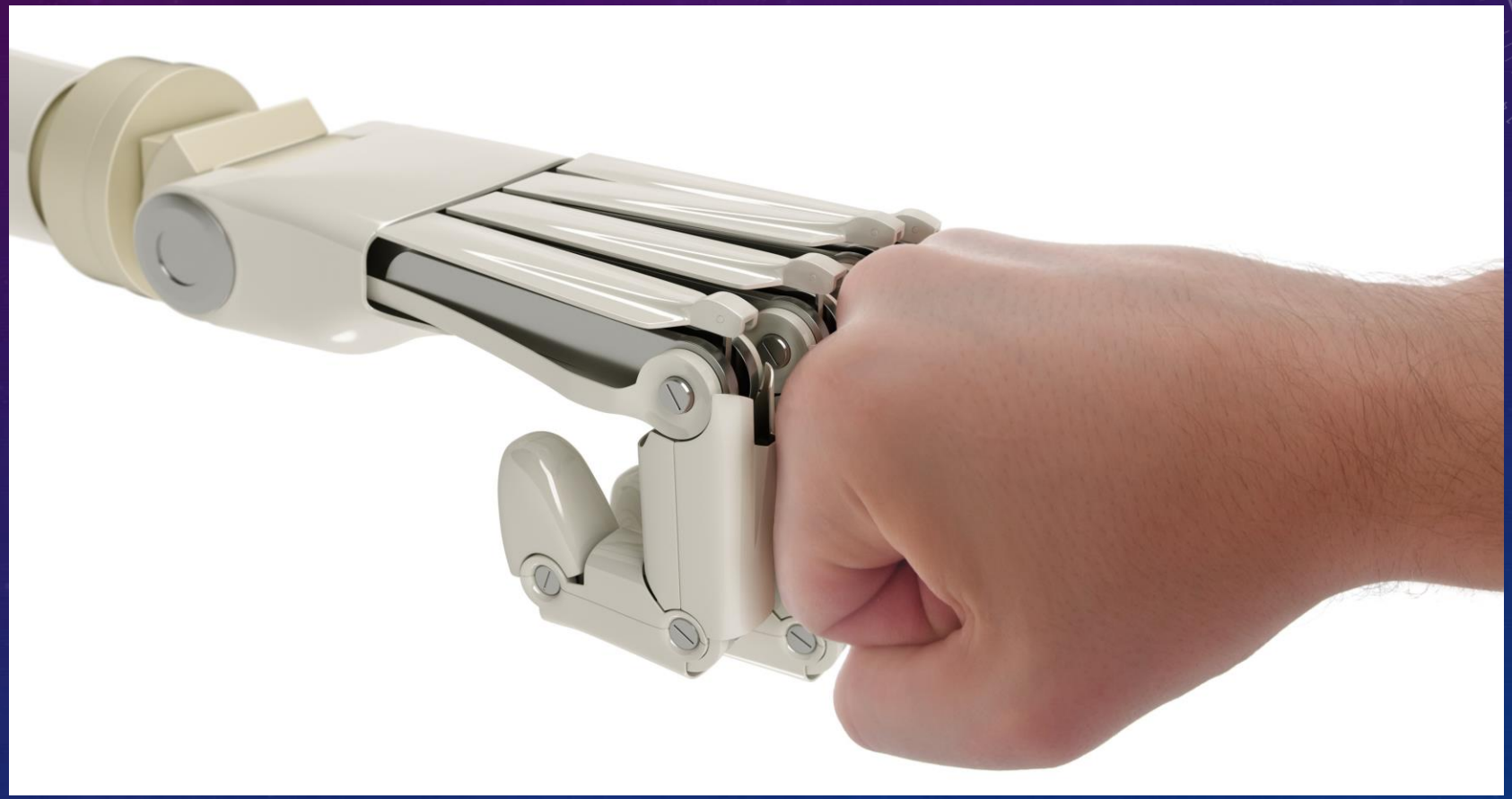



무한 반복
끝이 없네...
@tc
imgflip.com

REASON #1: 놀려고

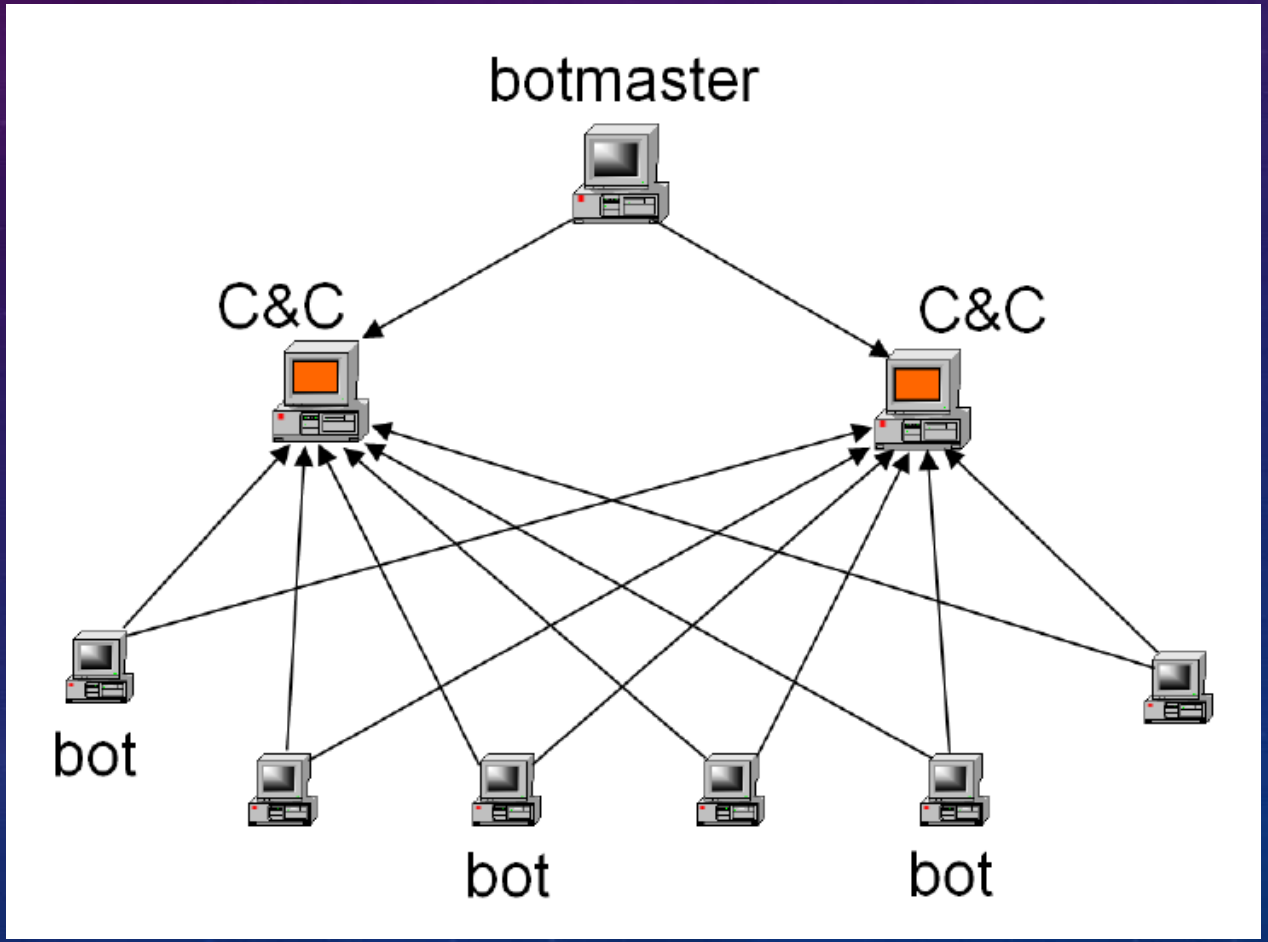
악성코드 자동분석





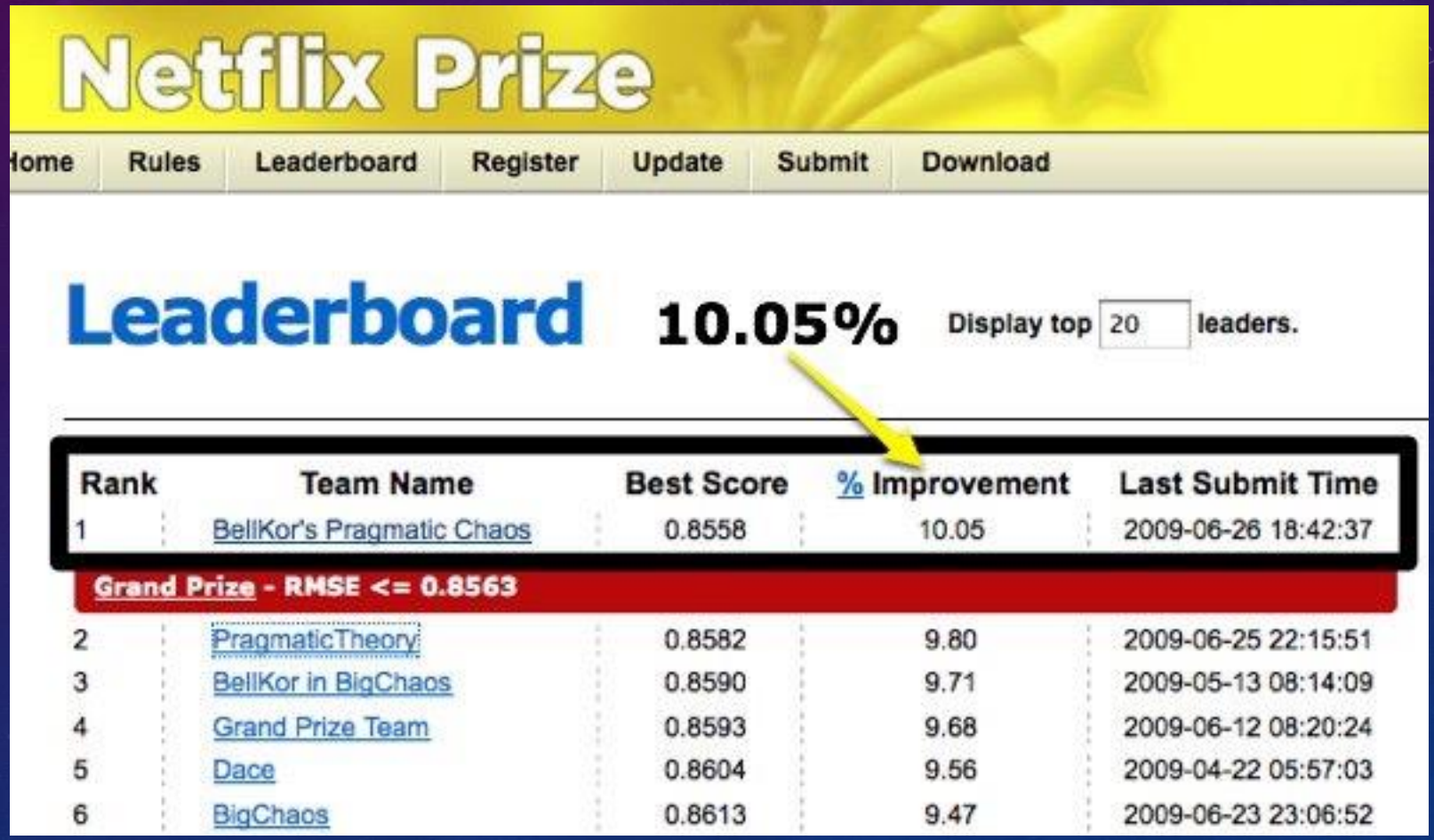
놀기도 지겹네.. 심심해...

LET'S MAKE BOTNET INTO ANALYSIS CLUSTER




음... 너무 위험해..
INITECH
imgflip.com

NETFLIX COMPETITION 을 하고 싶어요.



Netflix Prize

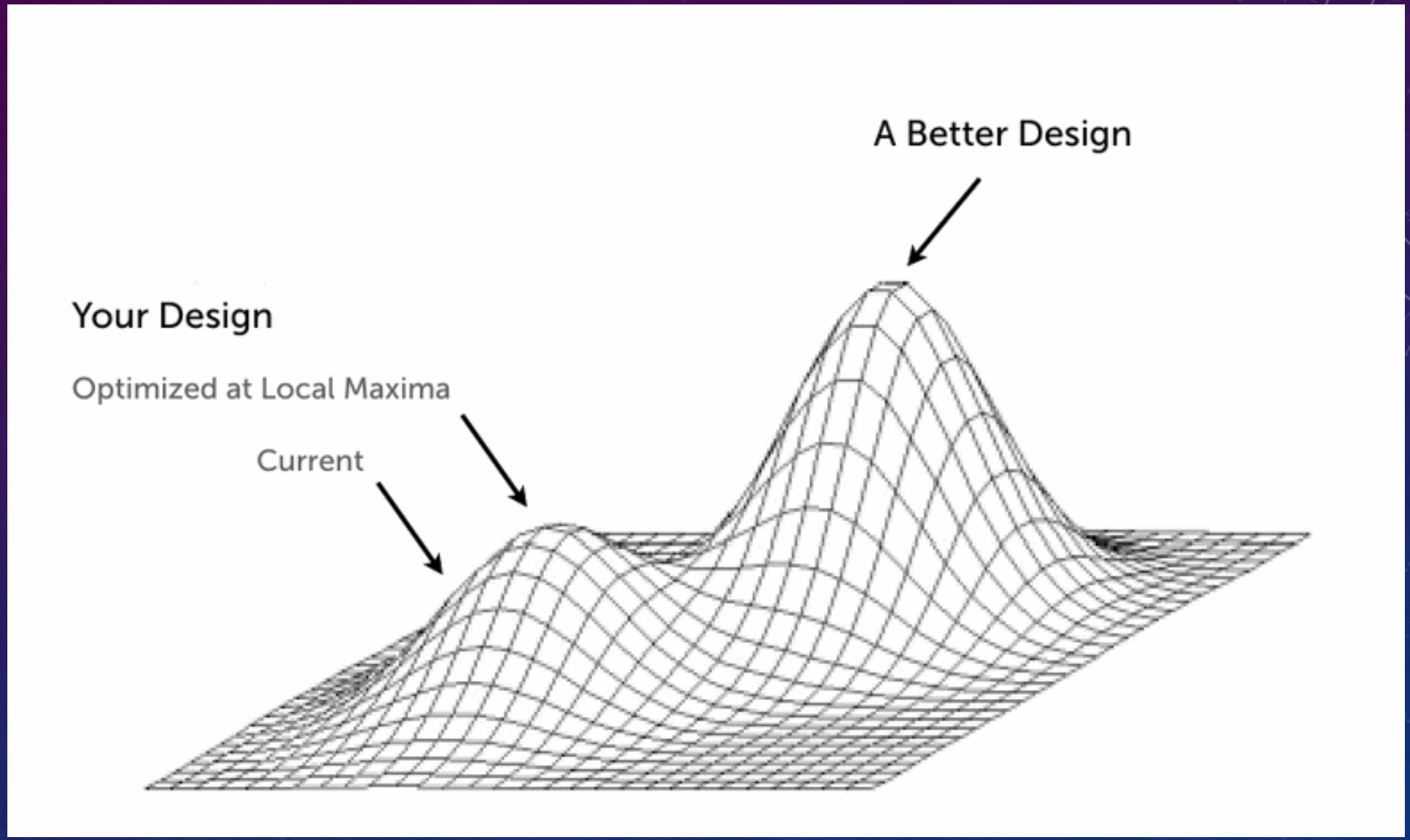
Home | Rules | Leaderboard | Register | Update | Submit | Download

Leaderboard 10.05% Display top 20 leaders.

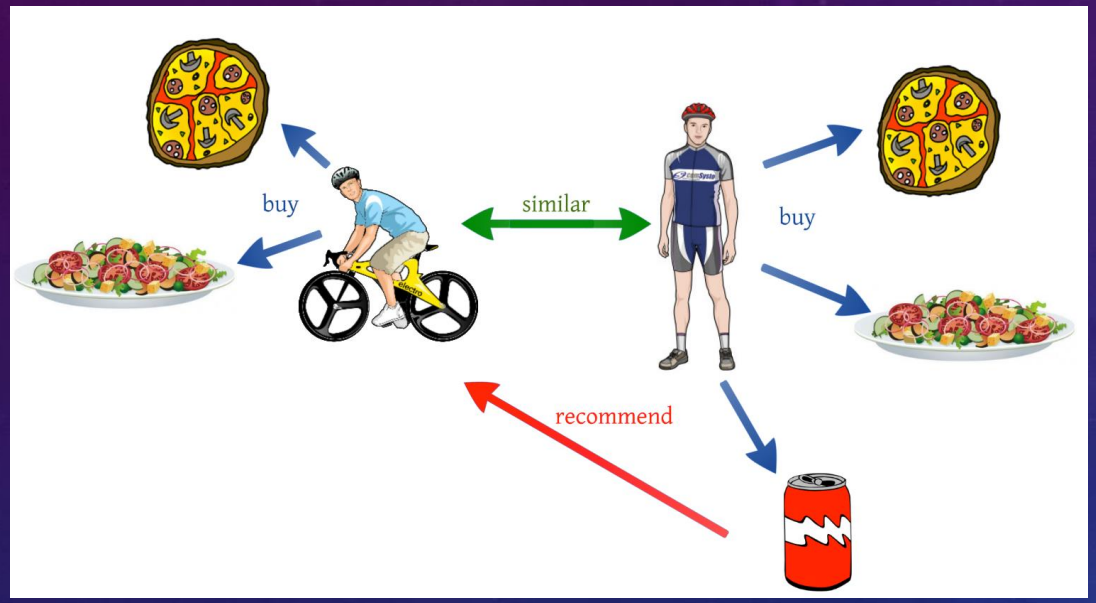
Rank	Team Name	Best Score	% Improvement	Last Submit Time
1	BellKor's Pragmatic Chaos	0.8558	10.05	2009-06-26 18:42:37
	Grand Prize - RMSE <= 0.8563			
2	PragmaticTheory	0.8582	9.80	2009-06-25 22:15:51
3	BellKor in BigChaos	0.8590	9.71	2009-05-13 08:14:09
4	Grand Prize Team	0.8593	9.68	2009-06-12 08:20:24
5	Dace	0.8604	9.56	2009-04-22 05:57:03
6	BigChaos	0.8613	9.47	2009-06-23 23:06:52


"MAYBE LATER"
memegenerator.net


SORRY
I QUIT
imgflip.com


A Better Design
Your Design
Optimized at Local Maxima
Current

스타텁



+

facebook

data

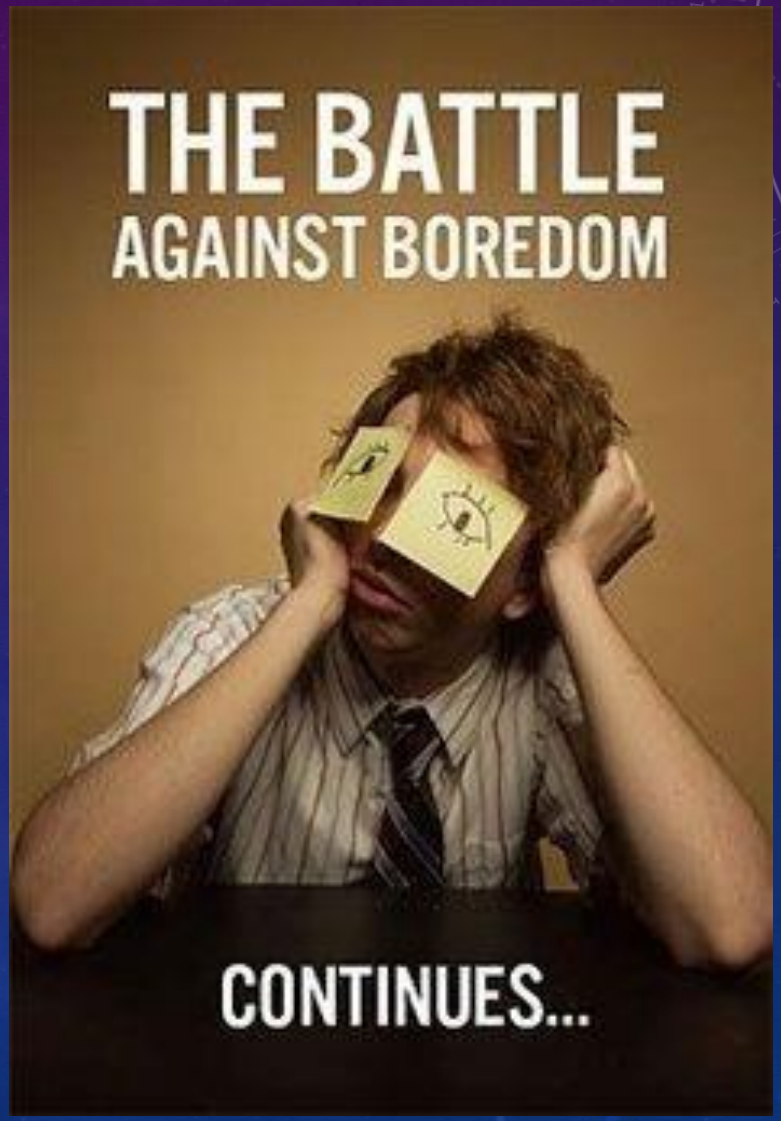
500+ Terabytes Per Day


THIS IS THE BEST
THING EVER


MEH
memegenerator.net

REASON #2: 놀고 있어서.

- (스타텁에서 배운 스킬은 대기업에서는 못쓴다.)
- Kaggle 이나 해볼까?



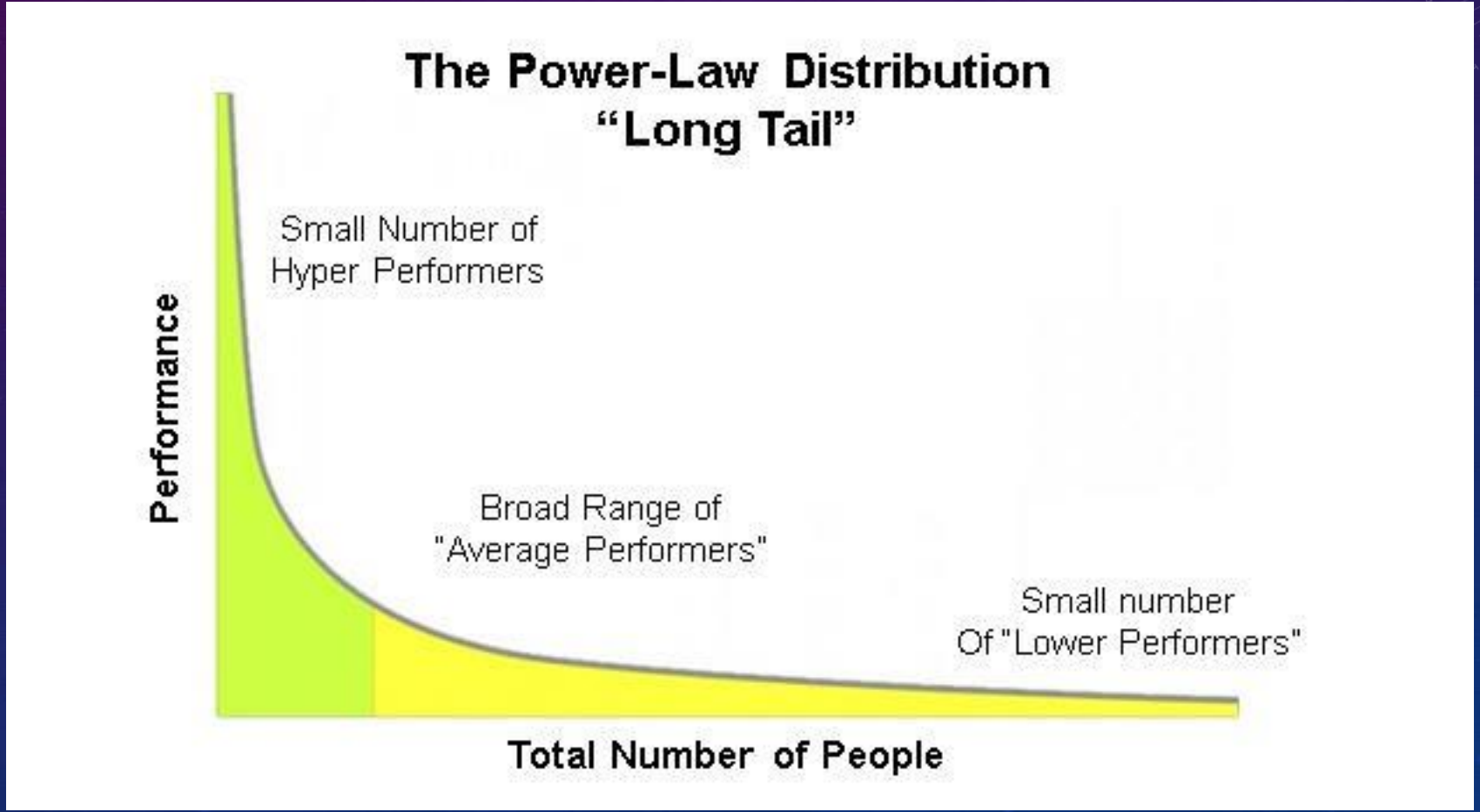
FIRST COMPETITION

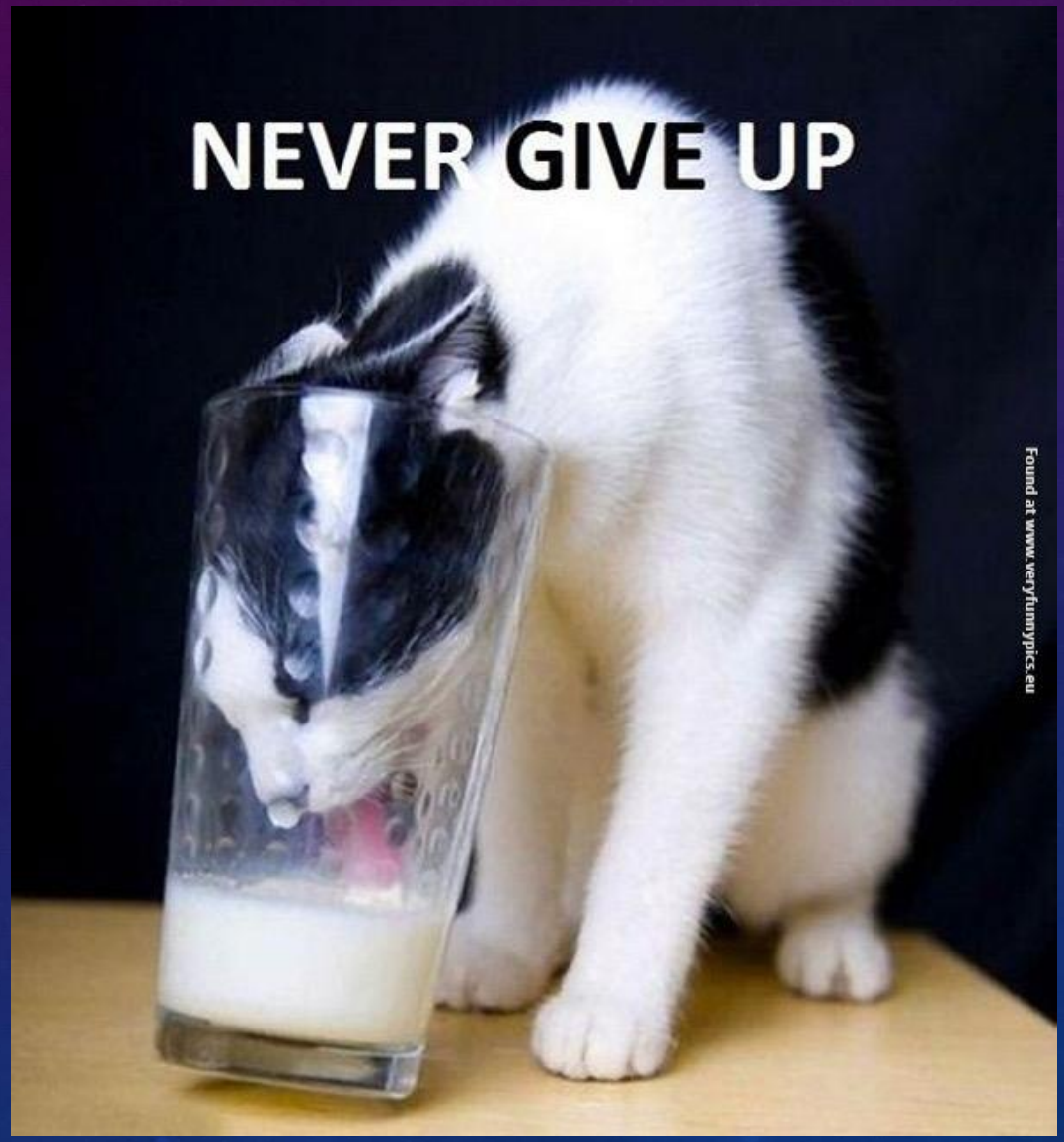

WEEK 1 OF 9
3RD PLACE
imgflip.com


WEEK 2 OF 9
10TH PLACE
imgflip.com


UFC
UFC
UFC

25 PERCENTILE 을 넘기지 못함

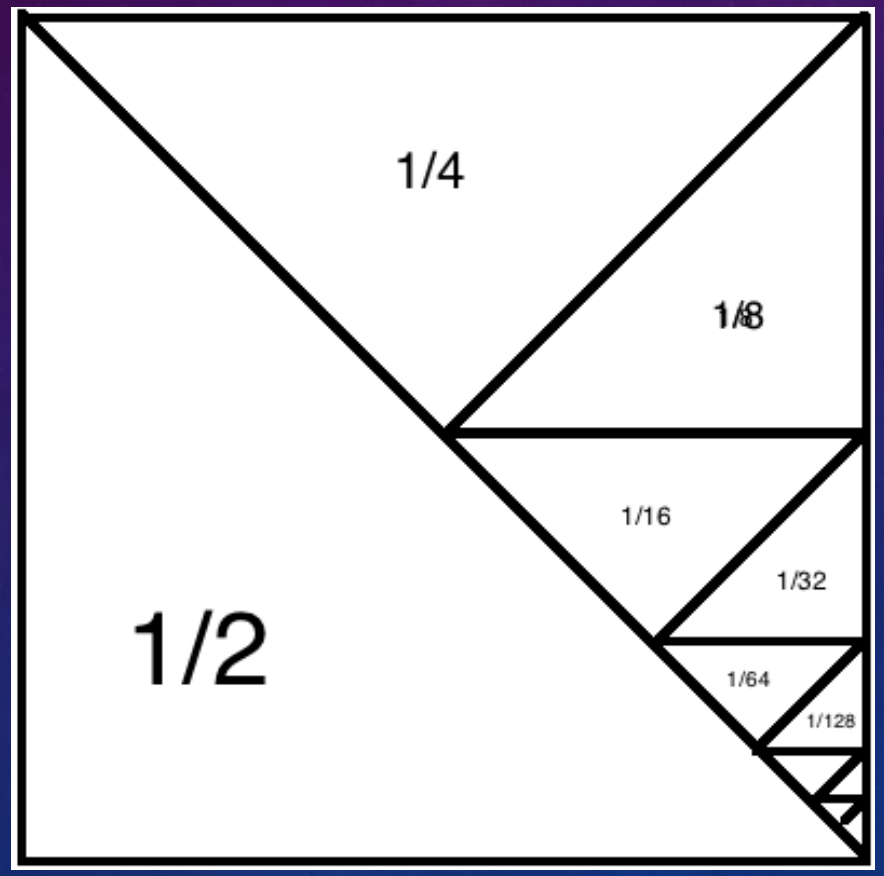



NEVER GIVE UP
Found at www.veryfunnypics.eu

KAGGLE FORUM

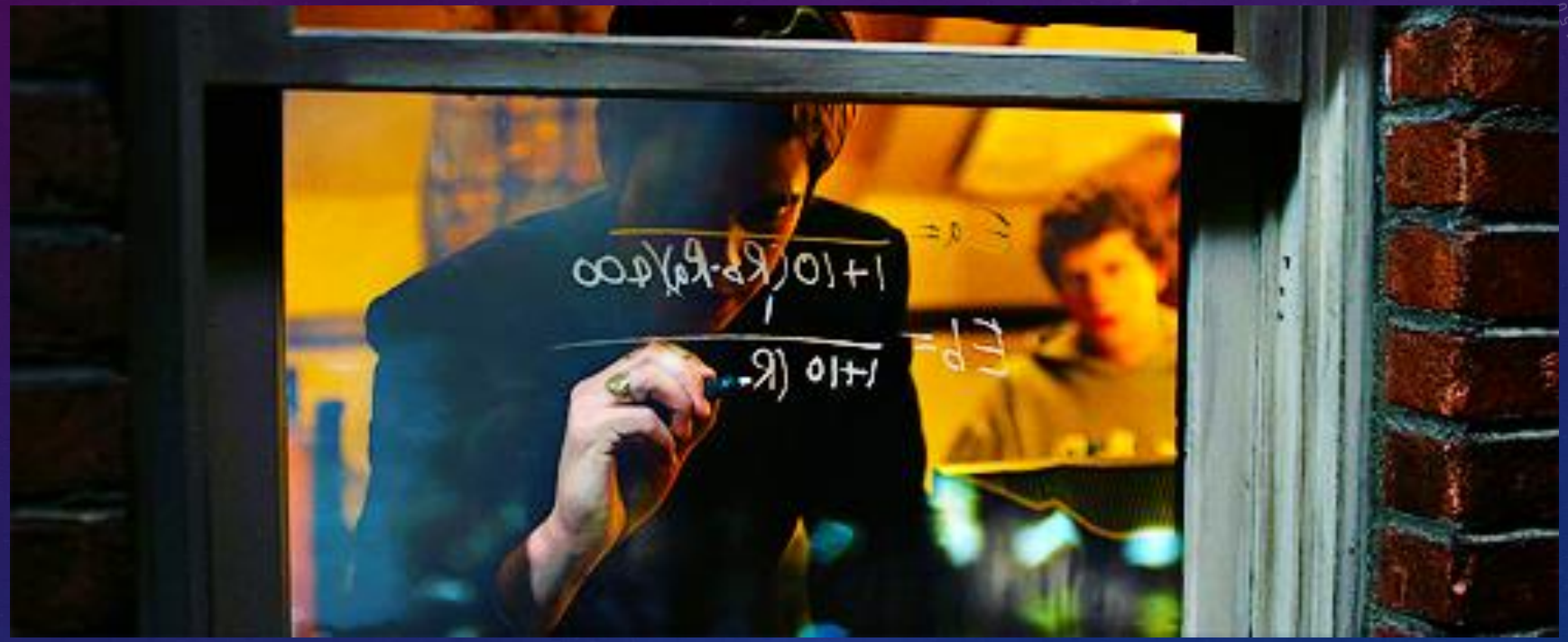
- competition 끝나면, 비밀 공개.
- 하나씩 배우기 시작함.
- 온간 잡 기술을 터득.

인터넷에 있는 기계학습 자료 다 읽기 시작함




ERROR METRIC 이 SQUARE니까..
PRE-PROCESSING 을 해서 NOISE를 골라서 날리고..
imgflip.com

SOCIAL NETWORK 영화를 보다가,



영화보다 멈춰놓고 공부함(그녀 미쳐감)

Not logged in

Article | Talk | Read | Edit | View history

WIKIPEDIA
The Free Encyclopedia

- Main page
- Contents
- Featured content
- Current events
- Random article
- Donate to Wikipedia
- Wikipedia store

Interaction
- Help
- About Wikipedia
- Community portal
- Recent changes
- Contact page

Tools

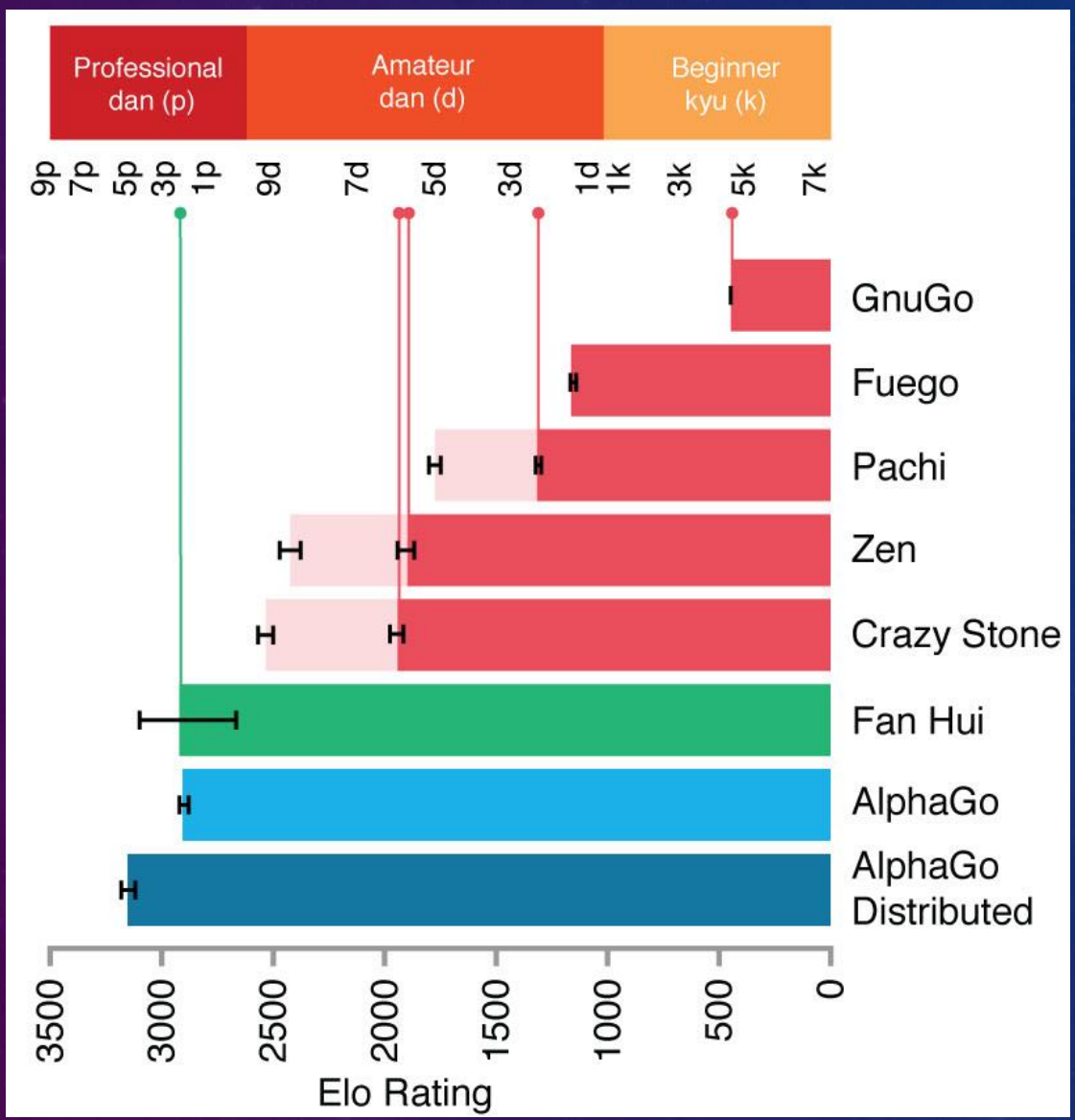
Elo rating system

From Wikipedia, the free encyclopedia

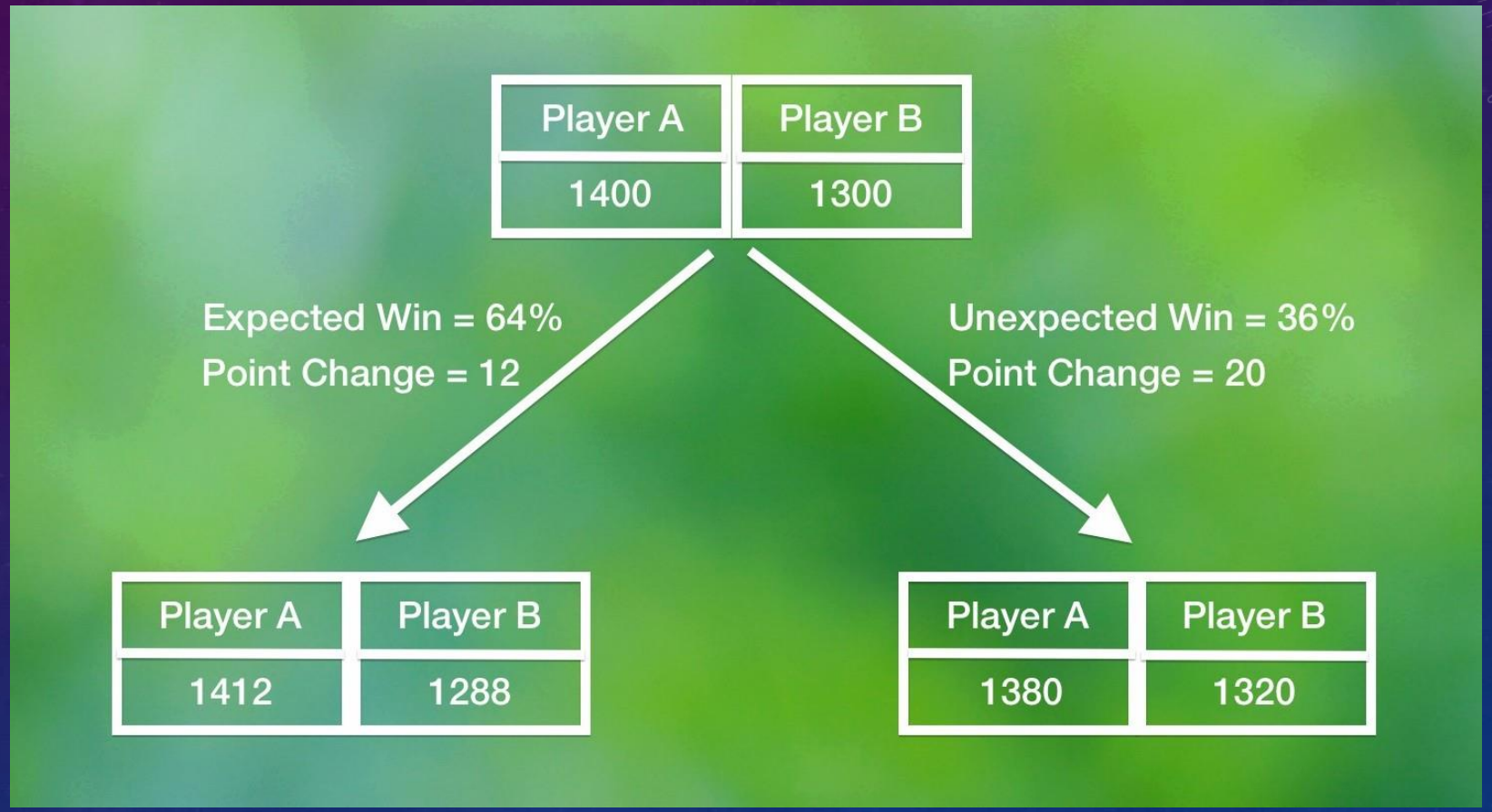
The **Elo rating system** is a method for calculating the relative skill levels of players in competitor-versus-competitor games such as chess. It is named after its creator Arpad Elo, a Hungarian-born American physics professor.

The Elo system was originally invented as an improved chess rating system but is also used as a rating system for multiplayer competition in a number of video games,[1] association football, American football, basketball,[2] Major League Baseball, *Scrabble*, snooker and other games.

The difference in the ratings between two players serves as a predictor of the outcome of a match. Two players with equal ratings who play against each other are expected to score an equal number of wins. A player whose rating is 100 points greater than their opponent's is expected to score 64%; if the difference is 200 points, then the expected score for the stronger player is 76%.


Professional dan (p)
Amateur dan (d)
Beginner kyu (k)
9p
7p
5p
3p
1p
9d
7d
5d
3d
1d
1k
3k
5k
7k
GnuGo
Fuego
Pachi
Zen
Crazy Stone
Fan Hui
AlphaGo
AlphaGo Distributed
3500
3000
2500
2000
1500
1000
500
0
Elo Rating

THE ALGORITHM 설명



24 HR HACKATHON, 4 HR LEFT TILL FINISH.

kaggle | Host | Competitions | Datasets | Kernels | Jobs | Community | John Park | Logout

Completed • $2,350 • 132 teams

Influencers in Social Networks

Sat 13 Apr 2013 – Sun 14 Apr 2013 (3 years ago)

Dashboard

- Home
  - Data
  - Make a submission
- Information
  - Description
  - Evaluation
  - Rules
  - Prizes
  - Partners
  - Preference Learning Tuto...
- Forum
- Leaderboard
  - Public
  - Private
- Visualization
- My Team
  - GitHub
- My Submissions

Competition Details » Get the Data » Make a submission

Predict which people are influential in a social network

The Big Data Hackathon

Data Science London and the **UK Windows Azure Users Group** in partnership with Microsoft and Peerindex, announce the ***Influencers in Social Networks*** competition as part of The Big Data Hackathon. This competition asks you to predict human judgements about who is more influential on social media.

The dataset, provided by **Peerindex,** comprises a standard, pair-wise preference learning task. Each datapoint describes two individuals, A and B. For each person, 11 pre-computed, non-negative numeric features based on twitter activity (such as volume of interactions, number of followers, etc) are provided.

SOCIAL RANKING COMPETITION

Leo Widrich curated a group on Public - switch to

Social Media Influencers

Tweet | Create Twitter list | Share

All starred accounts are estimated. If you know them, invite to join and their Peerl

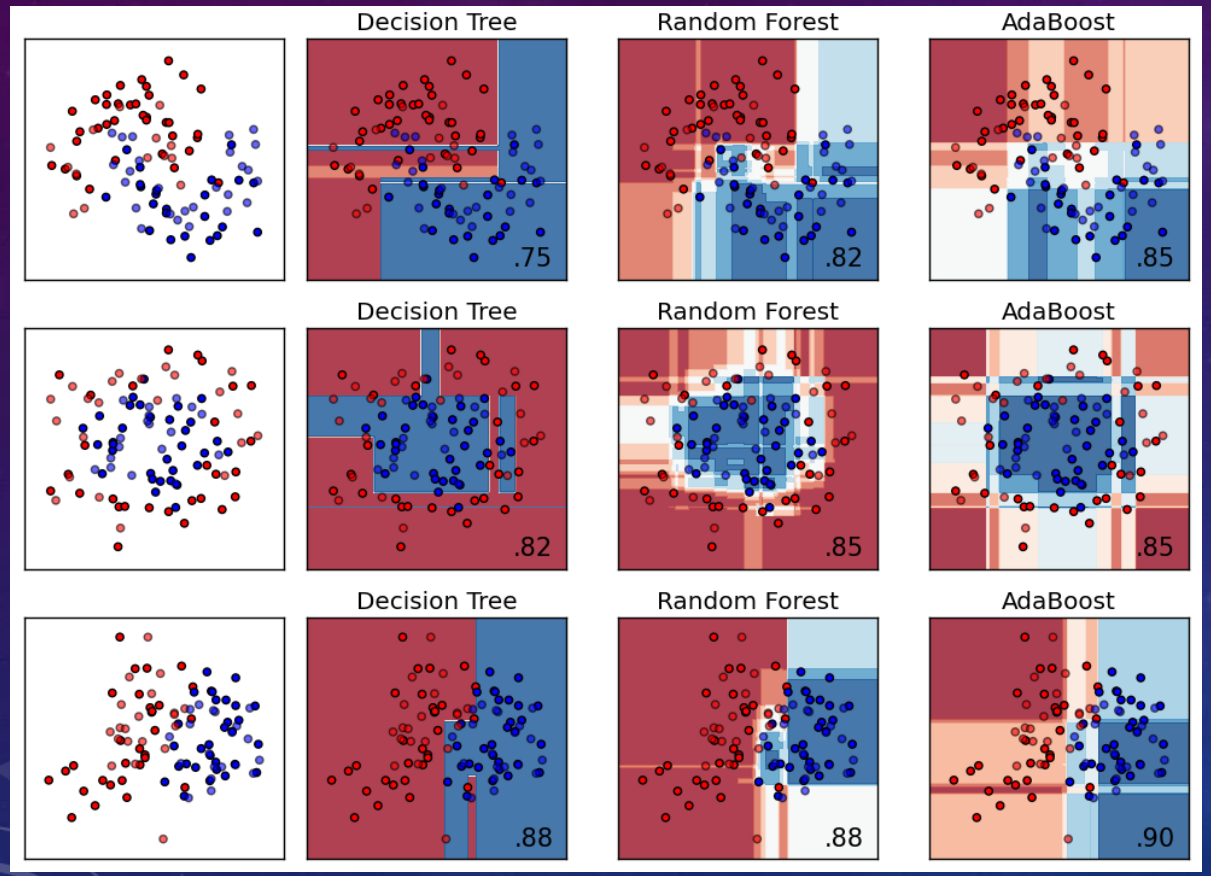
Avg. PeerIndex	59	Daily views	0	Total views	

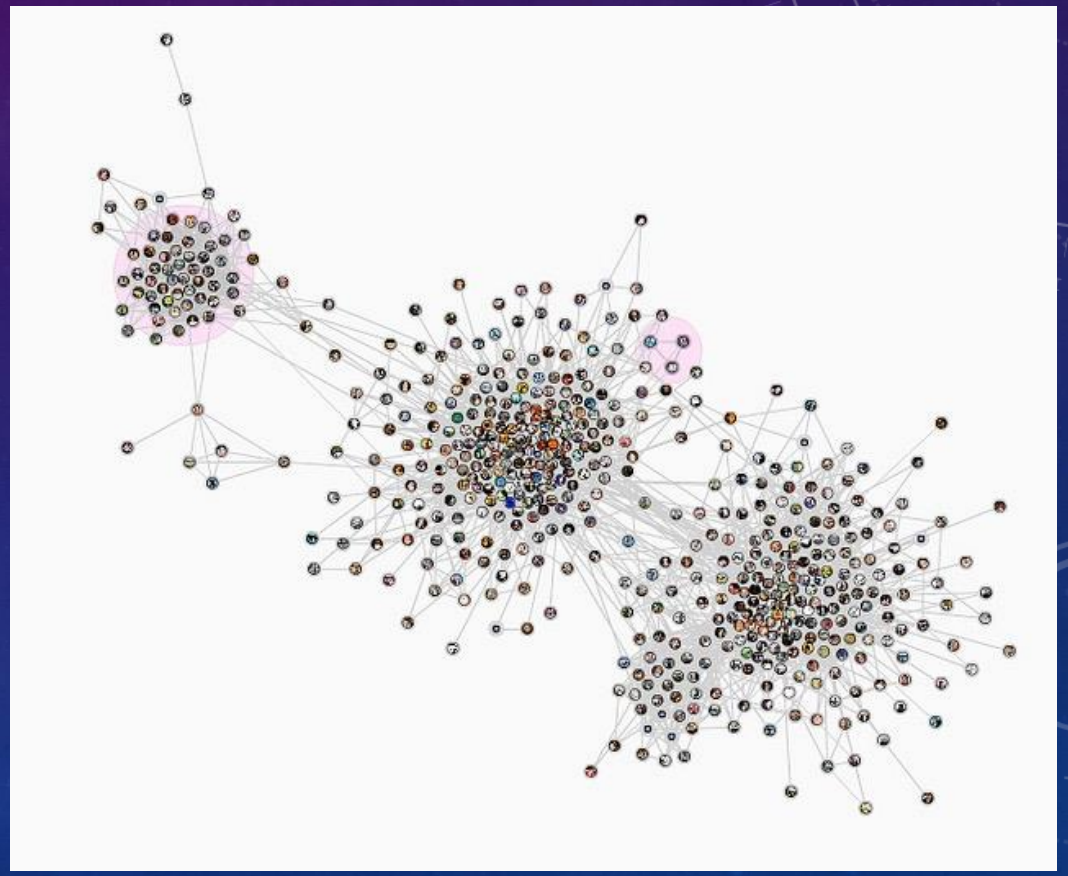
Add more users | Add collaborators (0) | Edit group details

	User	PeerIndex	Aut.	Act.	Aud.	Twitter
1.	Douglas Idugboe	79	74	87	95	@douglasi
2.	Michael Arrington	77	73	86	91	@arrington
3.	Marshall Kirkpatrick	74	69	87	87	@marshallk
4.	Alexia Tsotsis	74	70	86	86	@alexia

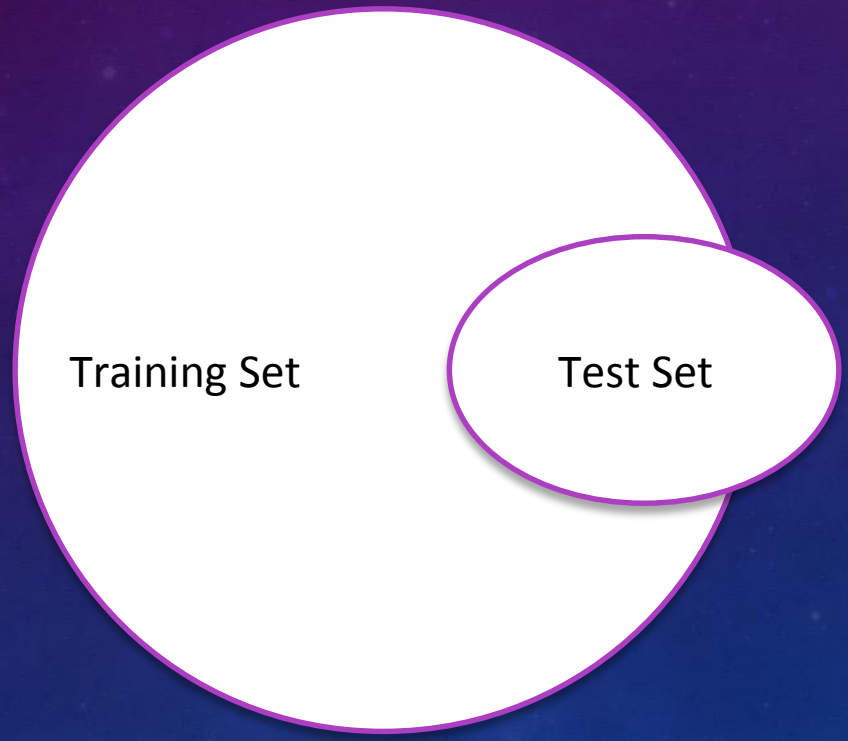
```
Choice,A_follower_count,A_following_count,A_listed_count,A_mentions_received,A_retweets_recei
0,228,302,3,0.583979387611,0.100503358535,0.100503358535,0.100503358535,0.362150068203,2,166.
0,21591,1179,228,90.4565058494,25.7982917525,5.70932864827,1.11115882218,5.17662002447,369,18
0,7310,1215,101,25.5036444818,9.55634744251,5.36151928351,0.591205606947,3.58971842047,95,68.
0,20,7,2,7.69082373179,0.277305640849,1.33150797971,0.100503358535,2.83062680131,6,2,96.16666
1,45589,862,2641,148.854279232,36.998883856,27.8817679309,3.33349180419,23.8612817152,551,127
0,285735,276251,3417,19.3275815086,7.29201558065,0.100503358535,0.100503358535,0.100503358535
0,285735,276251,3417,19.3275815086,7.29201558065,0.100503358535,0.100503358535,0.100503358535
1,9512,12,213,52.1670684554,23.1822884278,0.362150068203,0.100503358535,6.80604525921,195,11.
1,2273871,4524,11946,6782.40533805,2944.52424508,12.9557232877,2.79186097666,66.2062083002,21
0,182598,1402,3831,145.844909693,74.002957632,23.5491892963,0.100503358535,15.9532425551,567,
0,3200,3256,146,0.290192078012,0.100503358535,0.362150068203,0.100503358535,0.362150068203,1,
0,3914,1439,165,2.4042122036,1.16194602972,0.100503358535,0.100503358535,0.100503358535,9,18.
1,23230,495,826,118.052431863,56.6682926506,6.37320537301,0.840719744654,9.82700494448,466,10
0,4321482,259,32741,7068.5940808,3773.24875881,2.10199402263,0.339522140564,10.2165316735,253
1,2088,419,148,4.25083325871,0.100503358535,0.100503358535,0.100503358535,0.100503358535,18,9
1,48711,22845,932,61.9248042388,18.5789010899,0.100503358535,0.100503358535,0.100503358535,22
0,4760,425,96,5.0119622628,1.40789589748,0.100503358535,0.100503358535,0.100503358535,22,3,27
0,15385,673,747,55.9935458813,22.3219449455,6.94623325492,0.341936471695,6.50397747795,202,15
1,265258,209,551,631.915946309,457.648549848,5.46098548773,0.100503358535,7.49812649824,2603,
1,723800,328,7257,1695.43019461,317.463664979,3.5810568194,1.85822627025,6.48408594887,5210,1
0,352126,499,4428,476.925046137,201.19920615,6.7102808036,0.100503358535,7.97454850173,1881,5
1,4947410,112,26319,1774.21392717,830.899475371,0.100503358535,0.100503358535,0.100503358535,
```

THE STANDARD SOLUTION

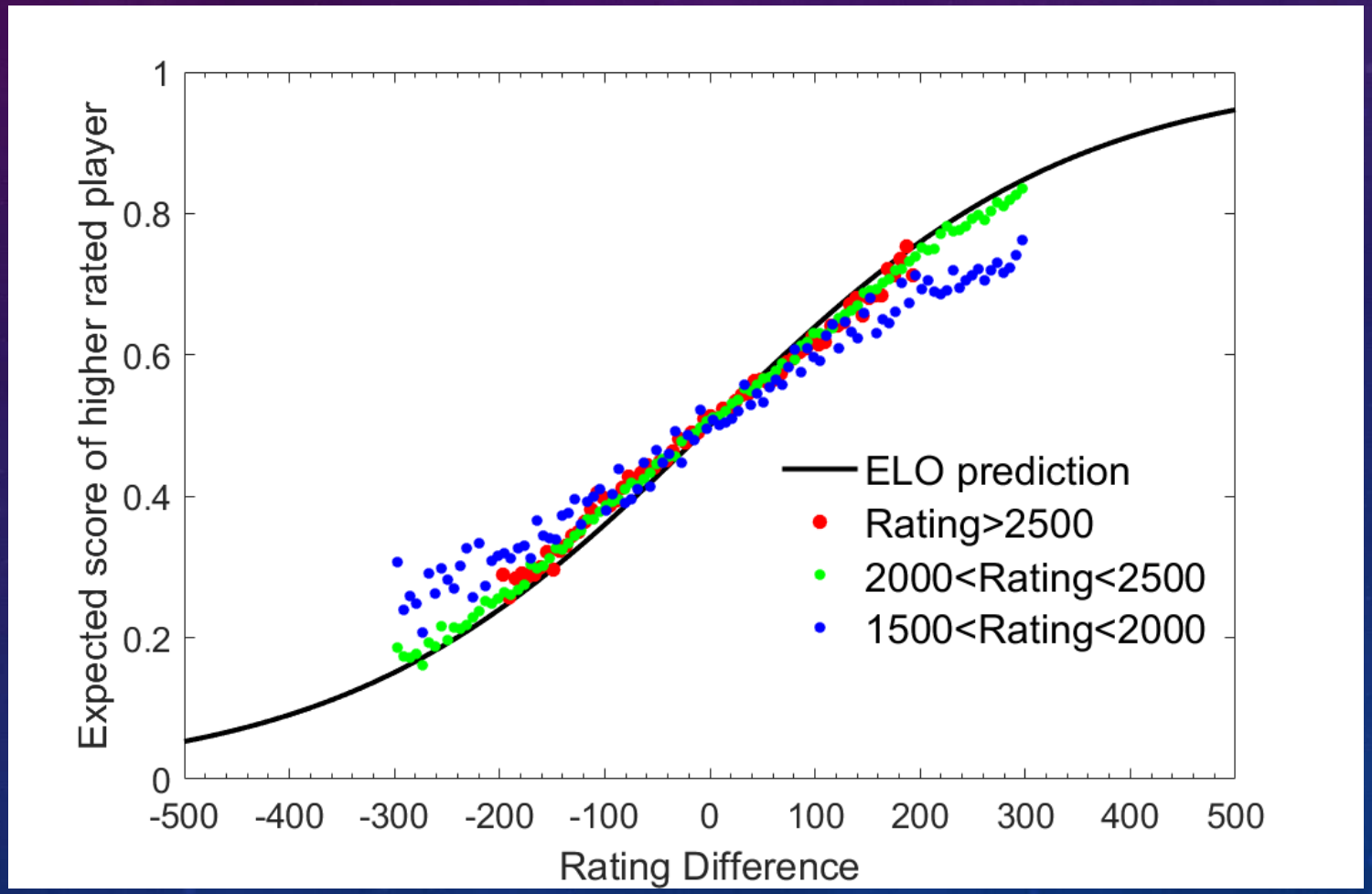




LARGE OVERLAP. NO NEED FOR “FEATURES”

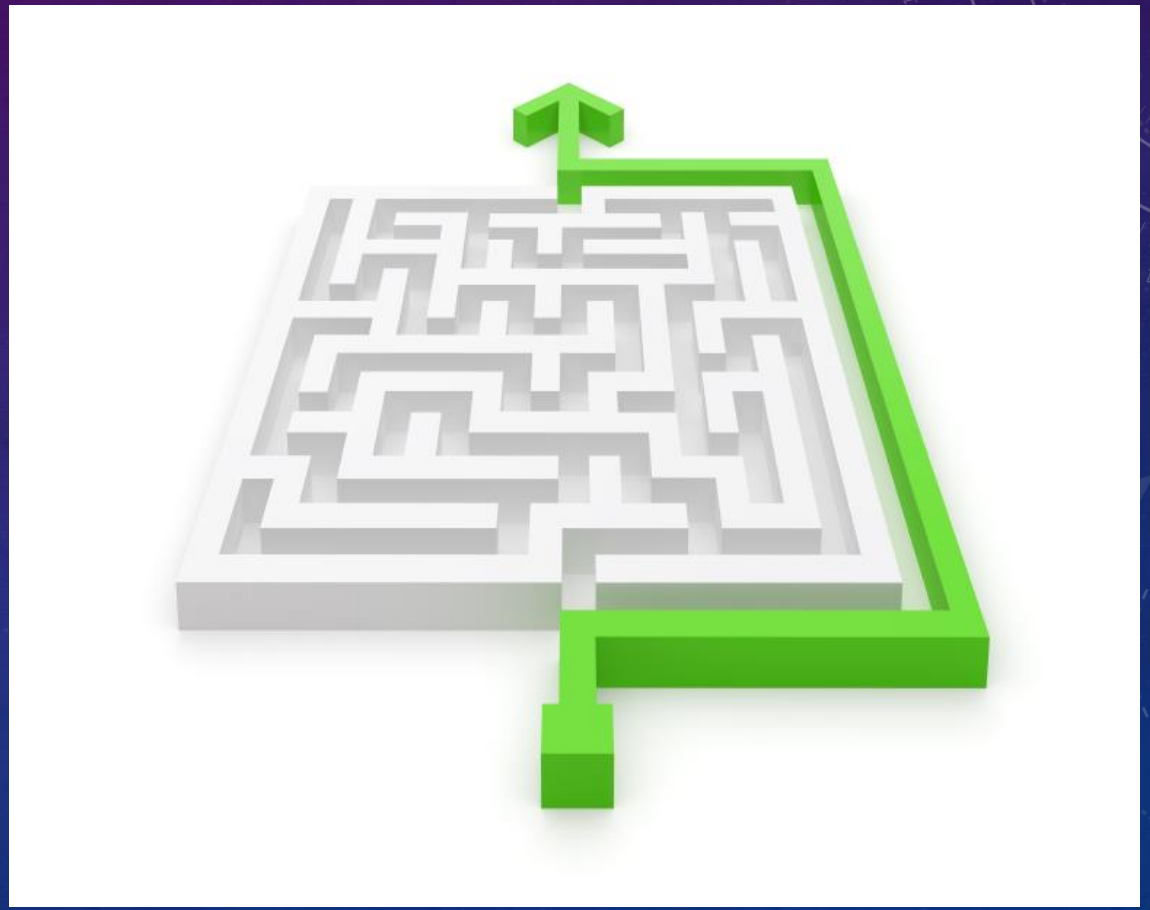


“HAIL MARY” MODEL.



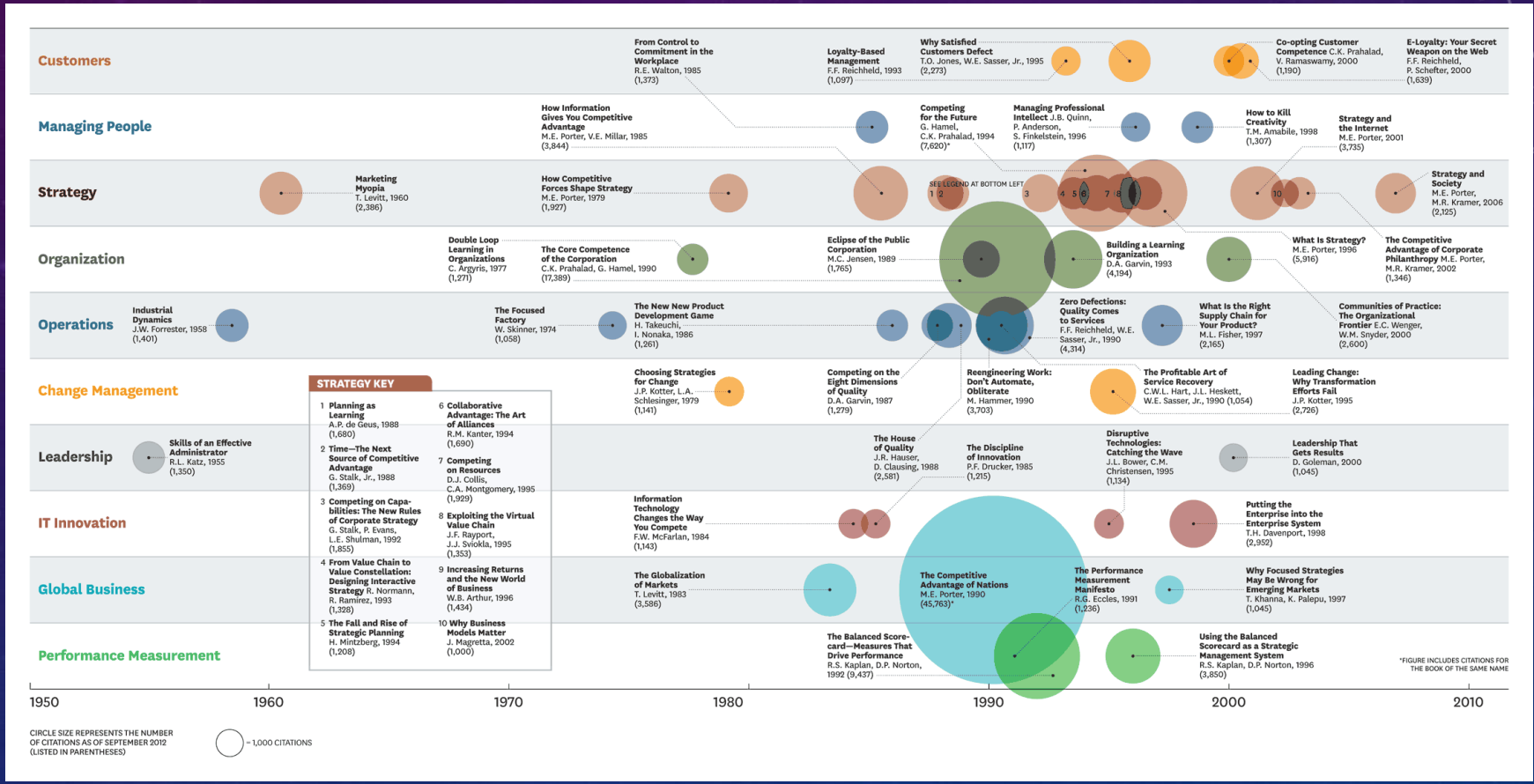
EPIPHANY (1)

- Top 10 rank.
- 20 lines of code



90YRS OF HARVARD BUSINESS REVIEW

THE STANDARD WAY



AUTO-GENERATED TITLES

- 3-gram max probability.
- Didn’t win.
- But, an honorable mention.

economics of money by national labor grievances

economics of us is the tests of rate

economics for your staffs informal networks culture

management women: debating the economics of things betterletter

economics of next?

beyond keynes: demand-side economics of business: recent books

the computer services more rational economics

myth of accounting review: a new economics

bottom-up economics and war policy

industrial advertising and how the new economics

your profits for review of economics

have to the new economics of buzzwords

ape in international economics of industrial pricing dilemmas

troubled legacy of the economics and diversity

economics and the public corporation: capital gains

economics relations with is the unpopular pay is

case of your next economics of manufacturing vision

economics of the press

EPIPHANY (2)

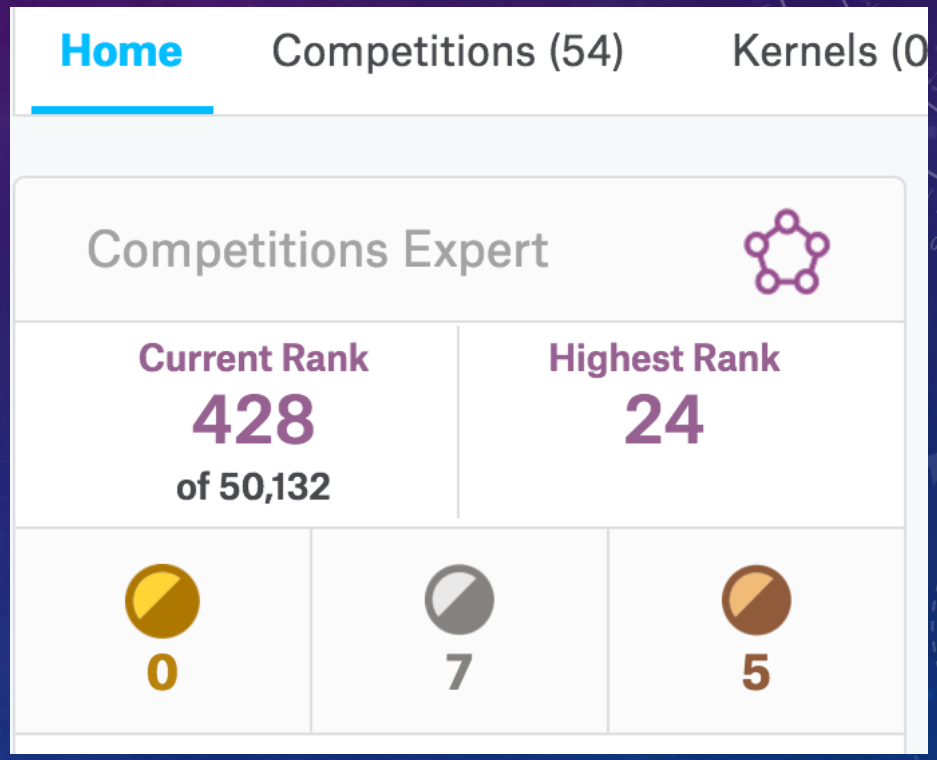
- It’s not magic.
- you just need the right “attack angle”

닥치는 대로 함.

- 6개월 동안, 밖에 안나감.
- 모든 사람과 연락 두절.

데이터가 보이기 시작함.

- still non-traditional.
- Peak 24th.
- more addicting than Starcraft.

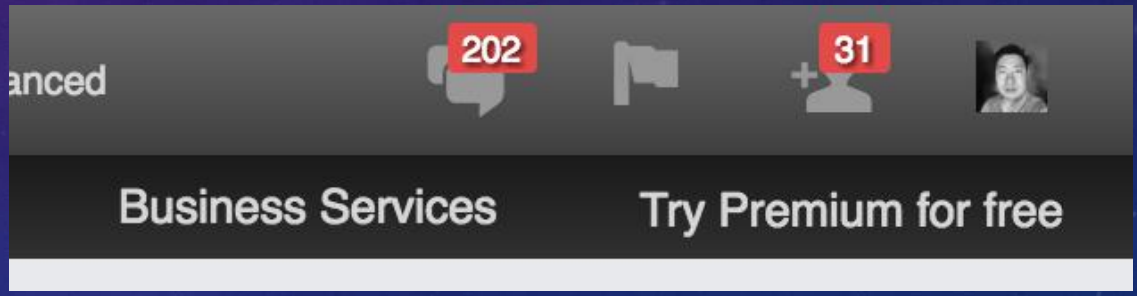


번 돈은 $300

- 50 competitions x 2 weeks = 100 weeks
- total rewards: $300
- $3/week.

WHAT HAPPENS AFTER

- Recruiters lined up.

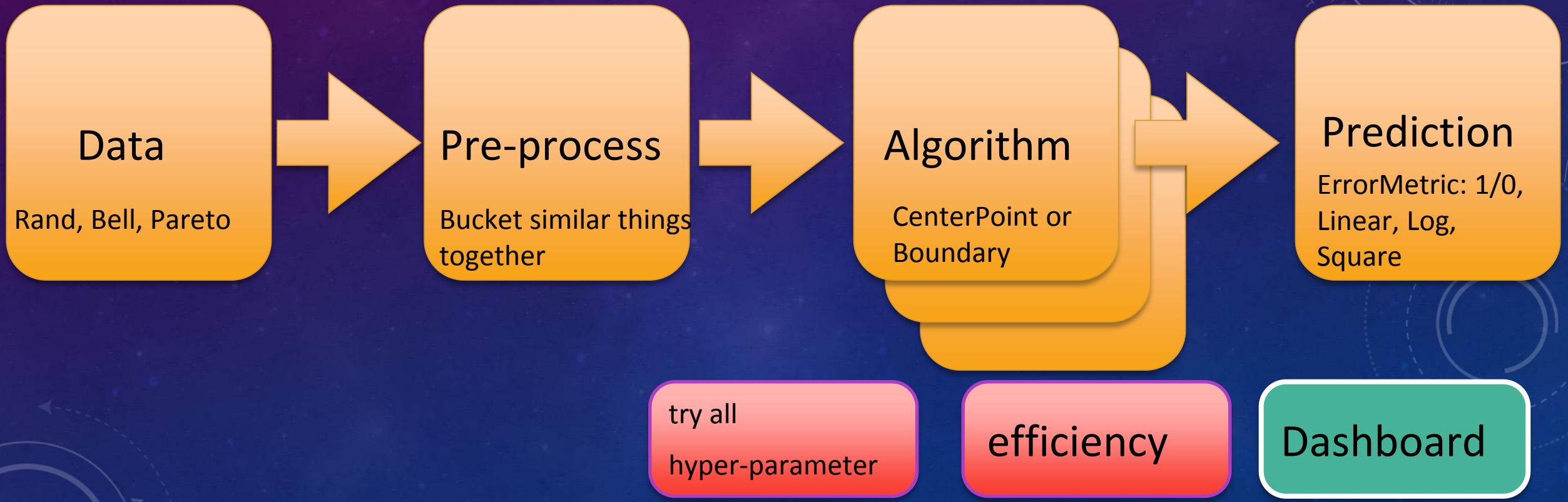


그럼면, 뿌려볼까나?

- 엑기스만 모아서.
- 정식은 아니지만, 실전에서는 잘 통함.

…. (5초 쉬고)

EVERYTHING PICTURE



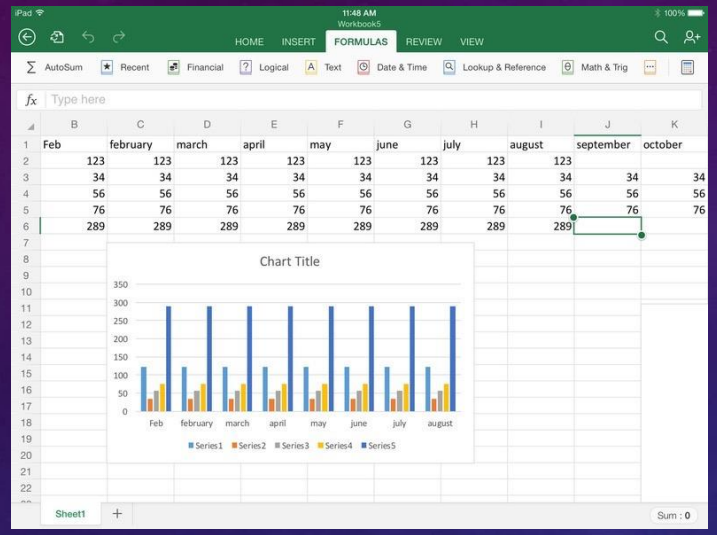
시작은... 2500 년전에

플라톤의 “동굴의 비유”


The fire
shadows cast
on wall
Prisoners
Roadway where
puppeteers perform

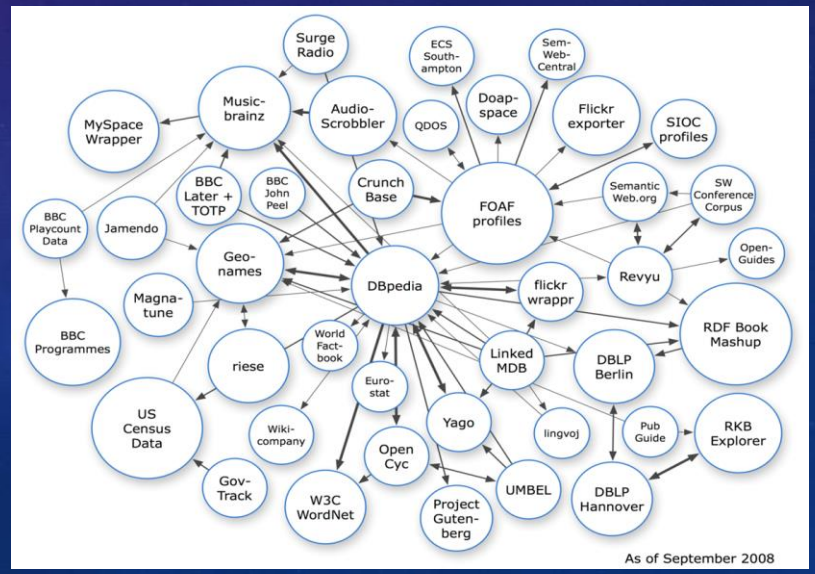
DATA (OBSERVATION)

4 TYPES OF DATA



Feb	february	march	april	may	june	july	august	september	october
123	123	123	123	123	123	123	123		
34	34	34	34	34	34	34	34	34	34
56	56	56	56	56	56	56	56	56	56
76	76	76	76	76	76	76	76	76	76
289	289	289	289	289	289	289	289		

Dies ist ein Blindtext. An ihm lässt sich vieles über die Schrift ablesen, in der er gesetzt ist. Auf den ersten Blick wird der Grauwert der Schriftfläche sichtbar. Dann kann man prüfen, wie gut die Schrift zu lesen ist und wie sie auf den Leser wirkt. Dies ist ein Blindtext. An ihm lässt sich vieles über die Schrift ablesen, in der er gesetzt ist. Auf den ersten Blick wird der Grauwert der Schriftfläche sichtbar. Dann kann man prüfen, wie gut die Schrift zu lesen ist und wie sie auf den Leser wirkt.

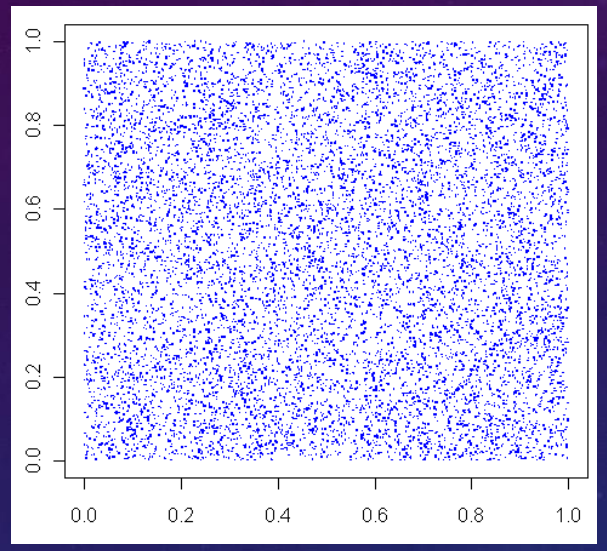


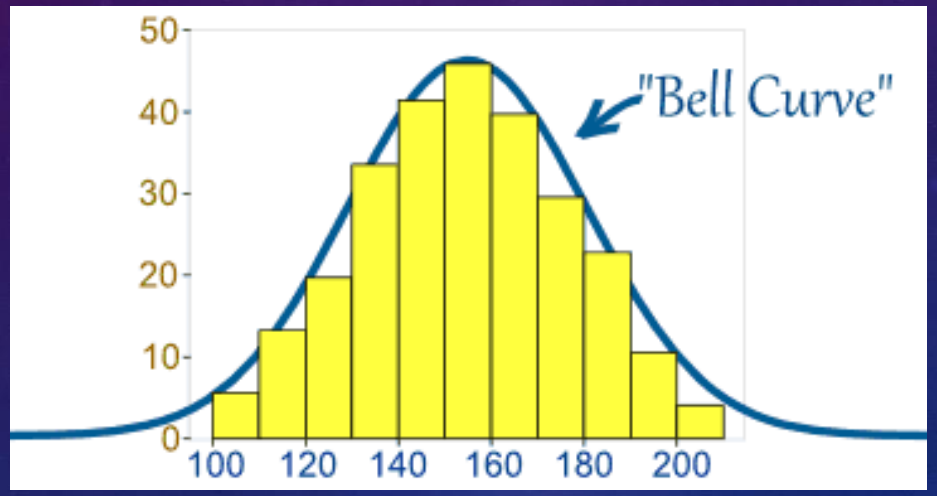
THE DATA FORMAT THAT ALGORITHM LIKES

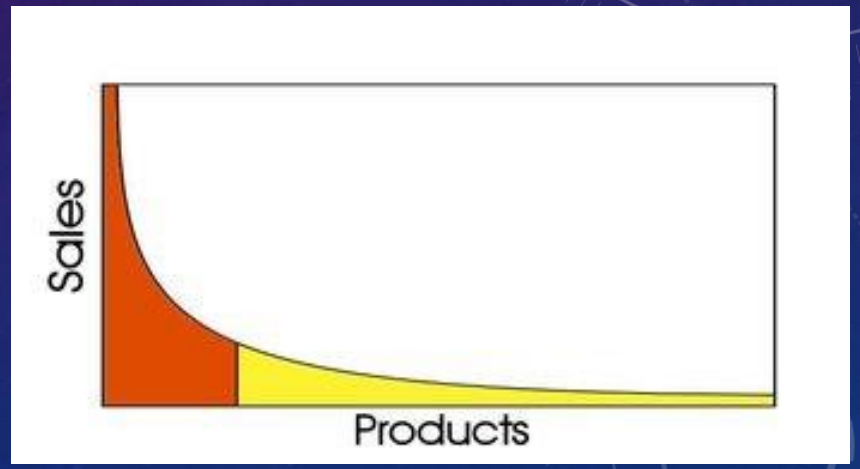
EventCode	Locality_in_words			State		LatDS		LatMS	LatSS	L
1	Mt Rumney	Tas	42	51	33	147	26	52	-42.8592	147.44
2	Dee Bridge	Tas	42	15	29	146	33	58	-42.2581	146.5661
3	Mt Wellington	Tas	42	54	00	147	14	00	-42.9	147.23
4	Hobart	Tas	42	52	58	147	19	49	-42.8828	147.3303
5	Wedge Bay	Tas	43	07	08	147	44	11	-43.1189	147.73
6	Forth Falls	Tas	41	23	16	146	12	13	-41.3878	146.2036
7	Cradle Mountain	Tas	41	38	07	145	56	41	-41.6353	145.94
8	Tunnack	Tas	42	27	12	147	27	42	-42.4533	147.4617
9	Western Creek	Tas	41	38	36	146	30	08	-41.6433	146.50
10	National Park	Tas	42	40	53	146	43	46	-42.6814	146.72
11	Weldborough	Tas	41	11	40	147	54	17	-41.1944	147.90
12	Lenah Valley	Tas	42	51	56	147	16	46	-42.8656	147.27
13	Cascades	Tas	42	53	46	147	17	35	-42.8961	147.2931
14	Lenah Valley	Tas	42	51	56	147	16	46	-42.8656	147.27
15	Domain - Hobart	Tas	42	51	44	147	19	21	-42.8622	147.32
16	Mt Wellington	Tas	42	54	00	147	14	00	-42.9	147.23
17	Mt Rufus	Tas	42	08	00	146	07	00	-42.1333	146.1167
18	Lake Leake	Tas	42	00	26	147	47	47	-42.0072	147.79
19	Spreyton	Tas	41	12	46	146	20	38	-41.2128	146.3439
20	Forth	Tas	41	11	23	146	15	00	-41.1897	146.25
21	Kallista	Tas	42	46	02	146	35	06	-42.7672	146.585
22	Snowy Mountains	Tas	42	56	44	146	41	42	-42.9456	

5km

DISTRIBUTIONS

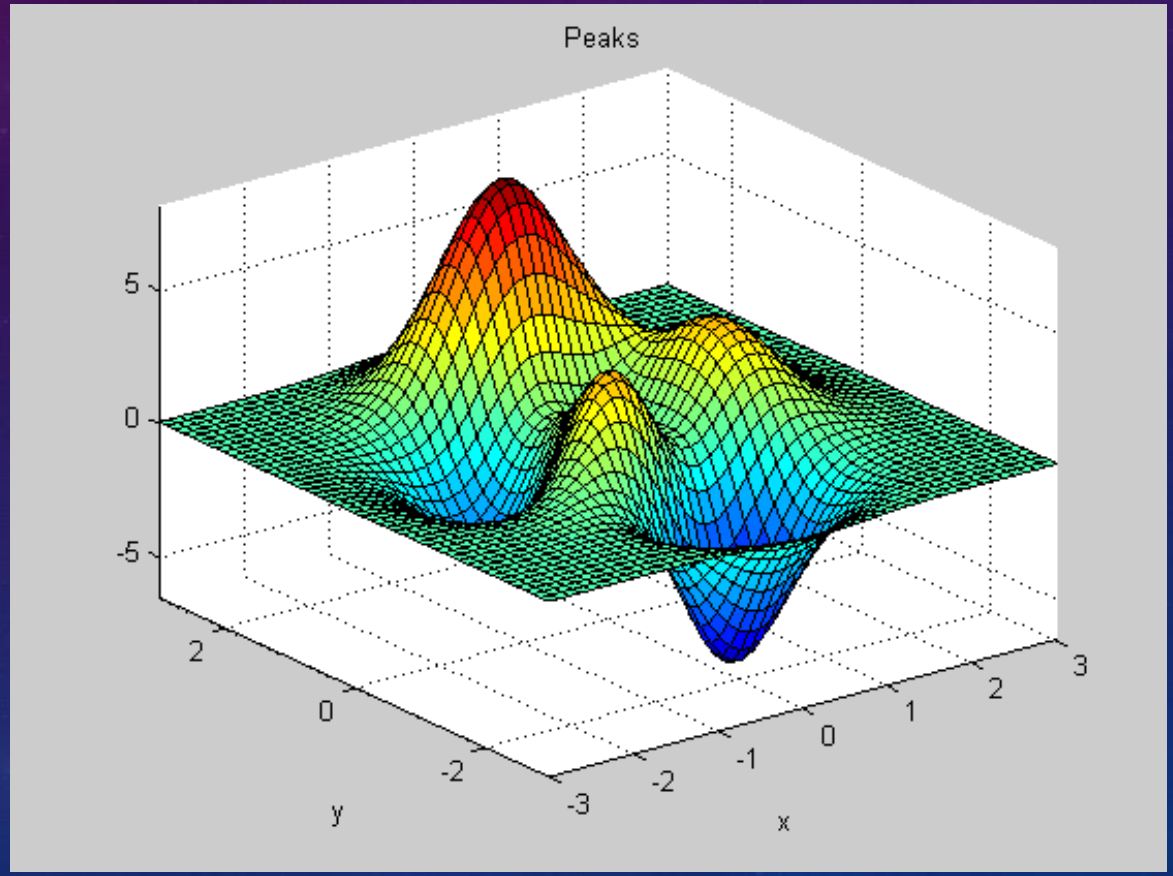






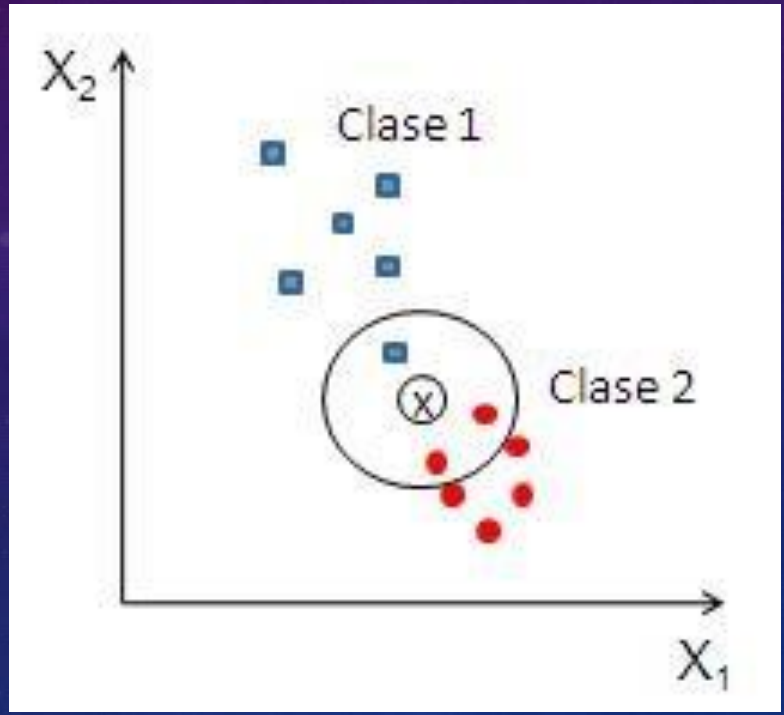
- random/no signal. OR condense/coherent/gaussian. OR wide/sparse.

SUPERVISED LEARNING

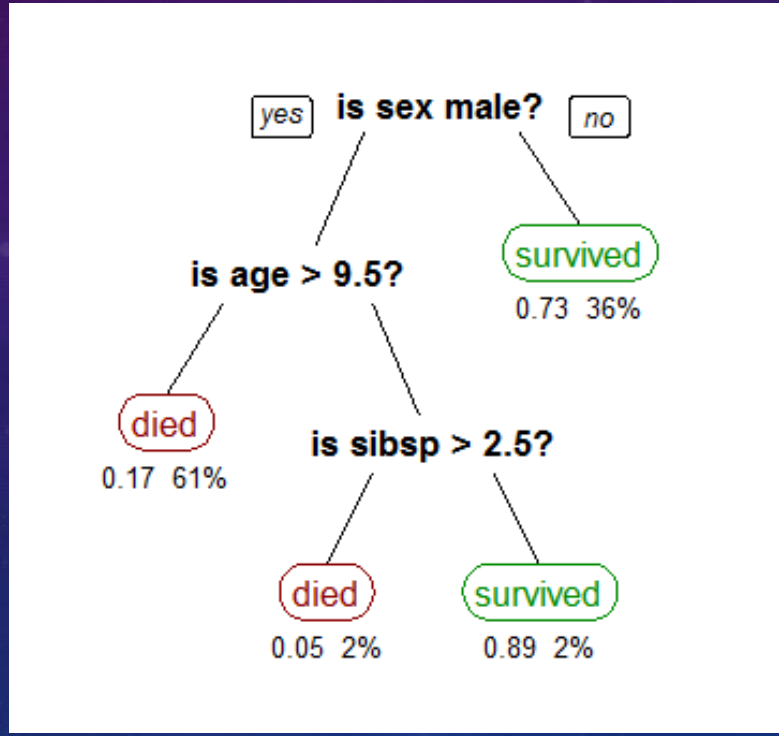


THAT’S ALL!

ALGORITHM (1) : K-NEAREST NEIGHBOR

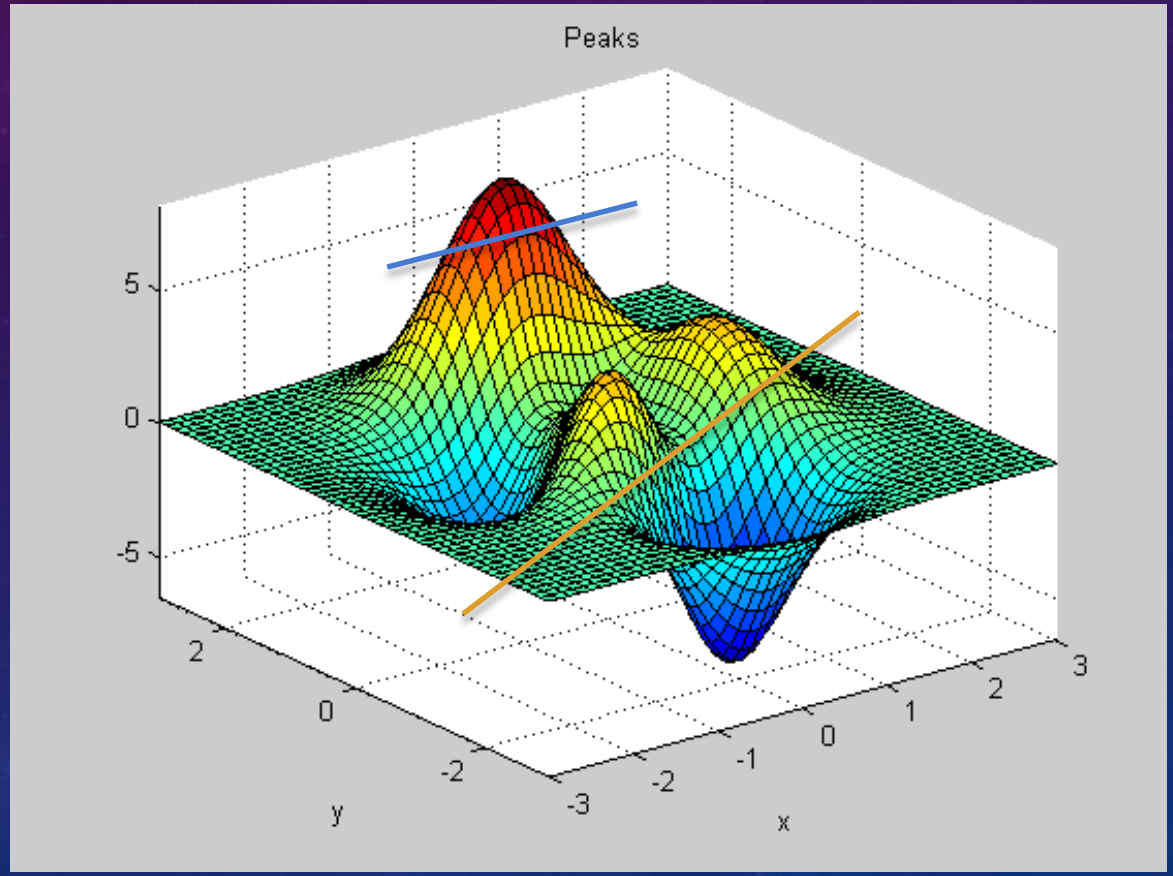


ALGORITHM (2): DECISION TREE



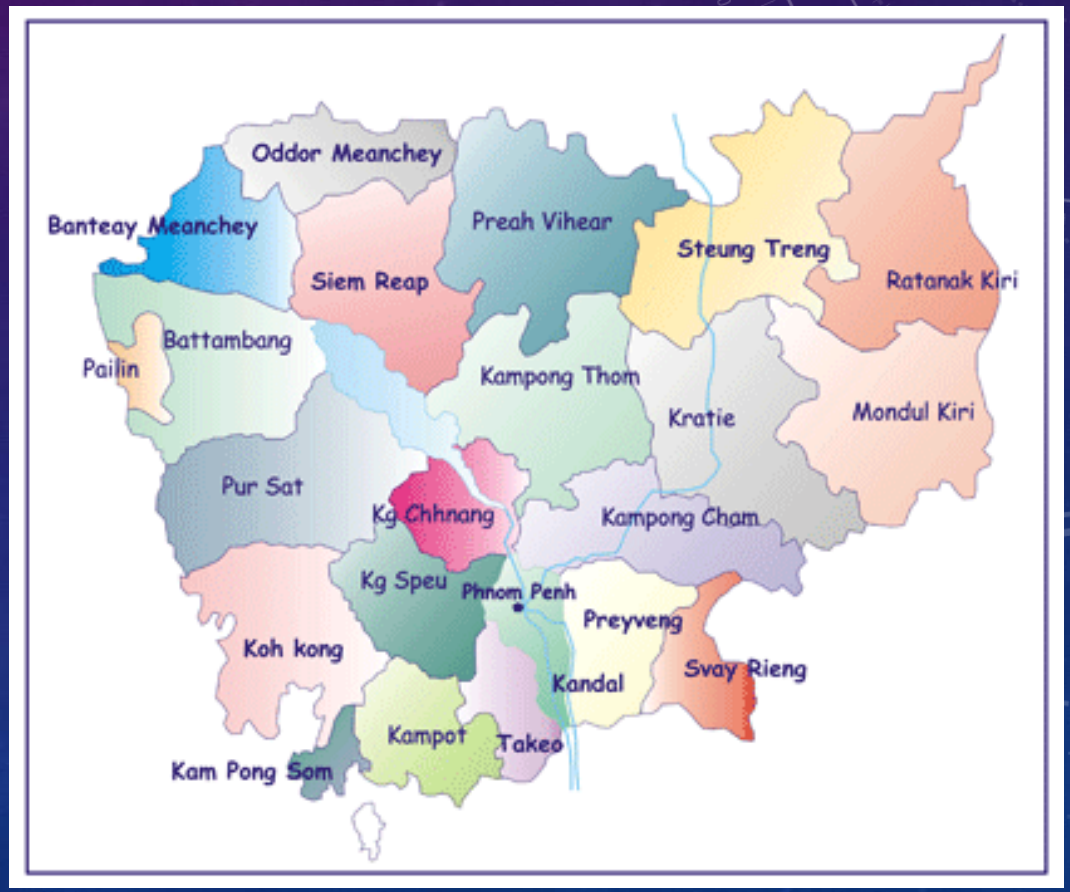
THAT’S ALL!

K-NN VS. DECISION TREE

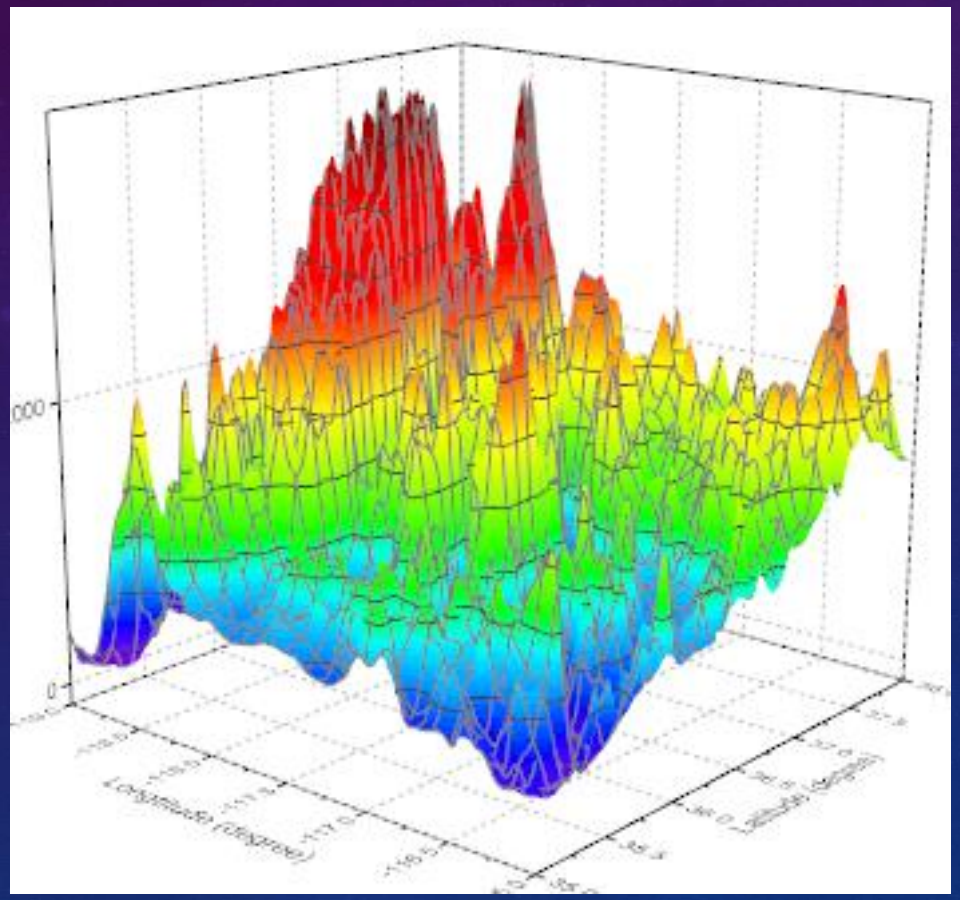


RECAP

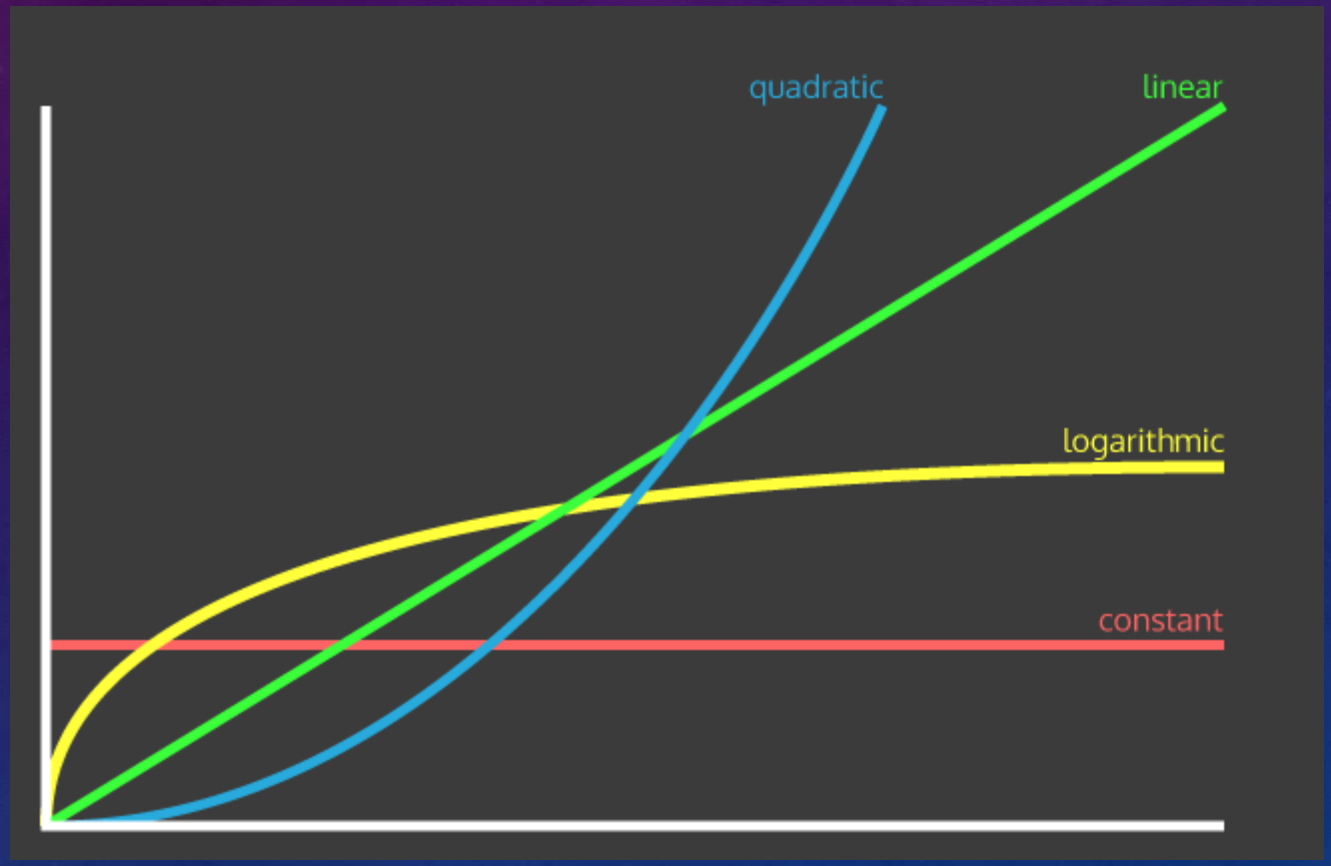

Thebes
Patrae
Marathon
Megara
Athens
Peiraeus
Sicyon
Salamis
Corinth
Elis
Phlius
Mycenae
Aegina
Olympia
Mantinea
Argos
Troezen
Messene
Sparta
Pylos
Methone
Gythium
Peloponnesus
www.mlahanas.de
Cythera


Oddor Meanchey
Banteay Meanchey
Preah Vihear
Steung Treng
Ratanak Kiri
Siem Reap
Battambang
Pailin
Kampong Thom
Kratie
Mondul Kiri
Pur Sat
Kg Chhnang
Kampong Cham
Kg Speu
Phnom Penh
Preyveng
Koh kong
Svay Rieng
Kandal
Kampot
Takeo
Kam Pong Som

TOO JAGGED = OVERFIT



ERROR METRIC




Data
Algorithm
Prediction

The features you use influence more than everything else the result. No algorithm alone, to my knowledge, can supplement the information gain given by correct feature engineering. (Luca Massaron)

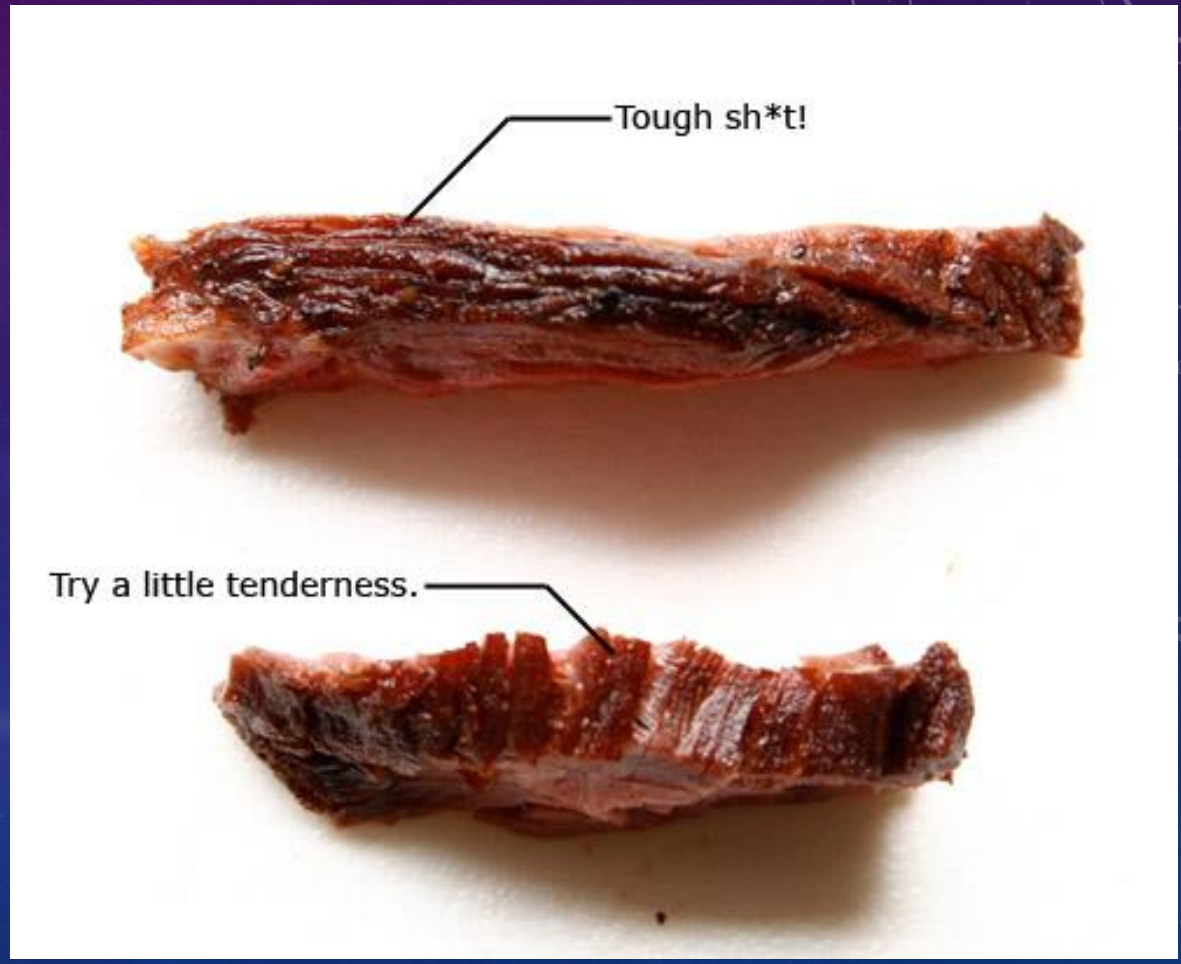
FEATURE ENGINEERING / UNSUPERVISED

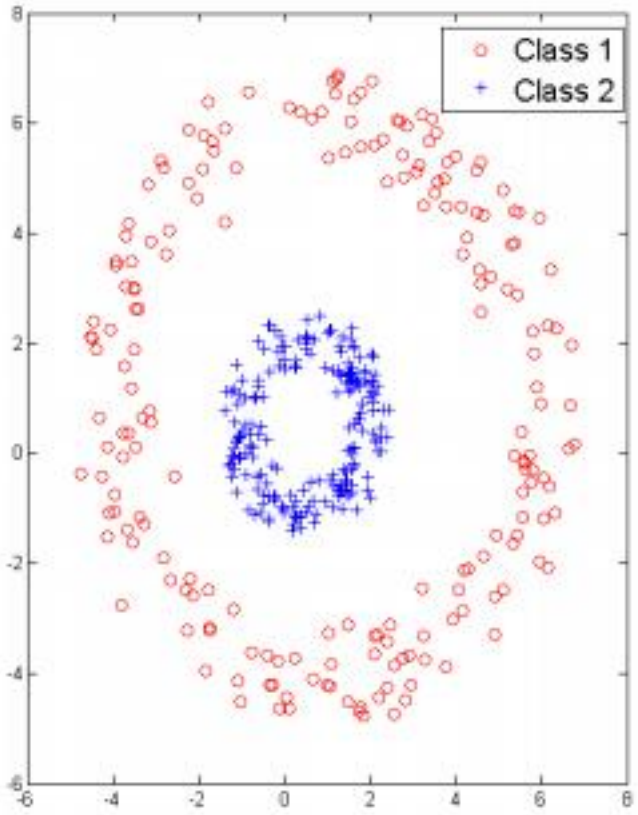
- Pre-processing
- Mathematical Formula
- Domain/Real-World Knowledge helps

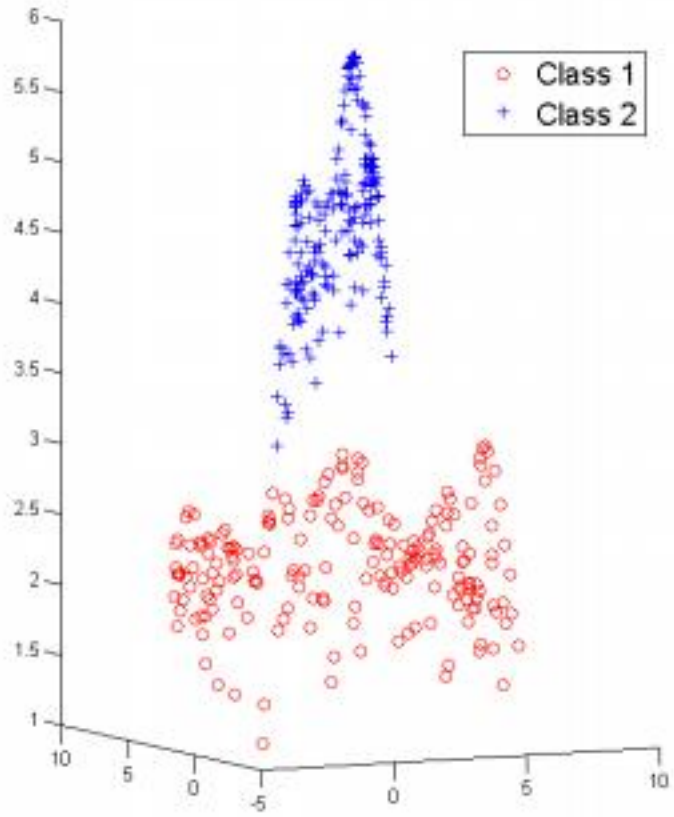
EASY TO BUCKET IT

NOT SO EASY TO BUCKET IT (REAL WORLD)

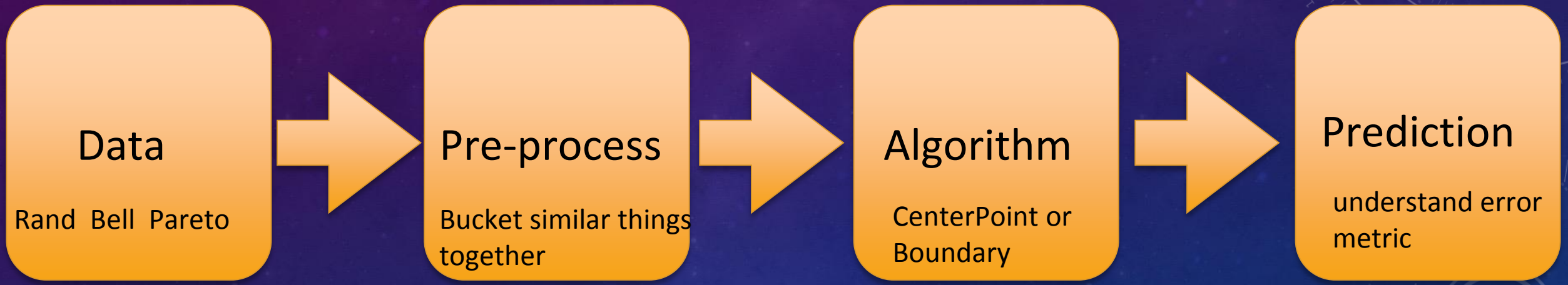
- go, went, going.
- start, starve, starbucks, startrek, starwars.



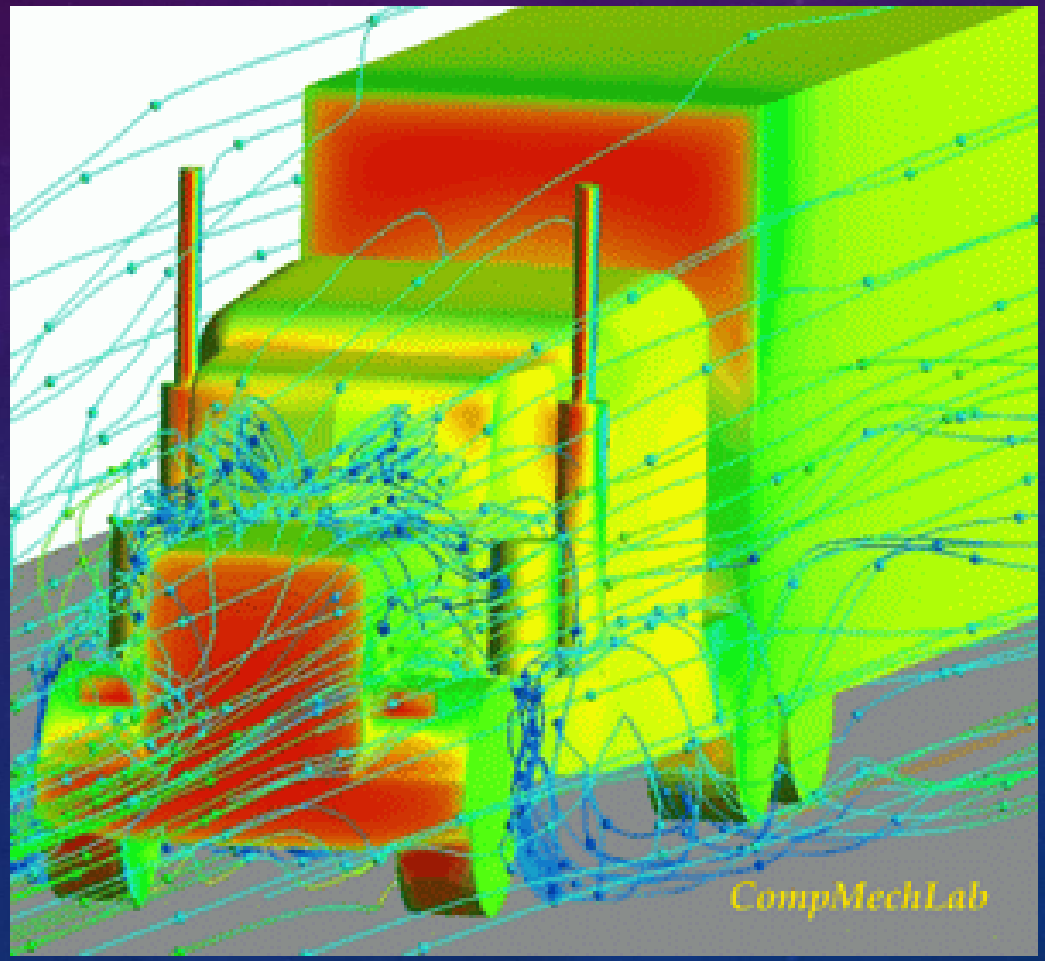

Class 1
Class 2

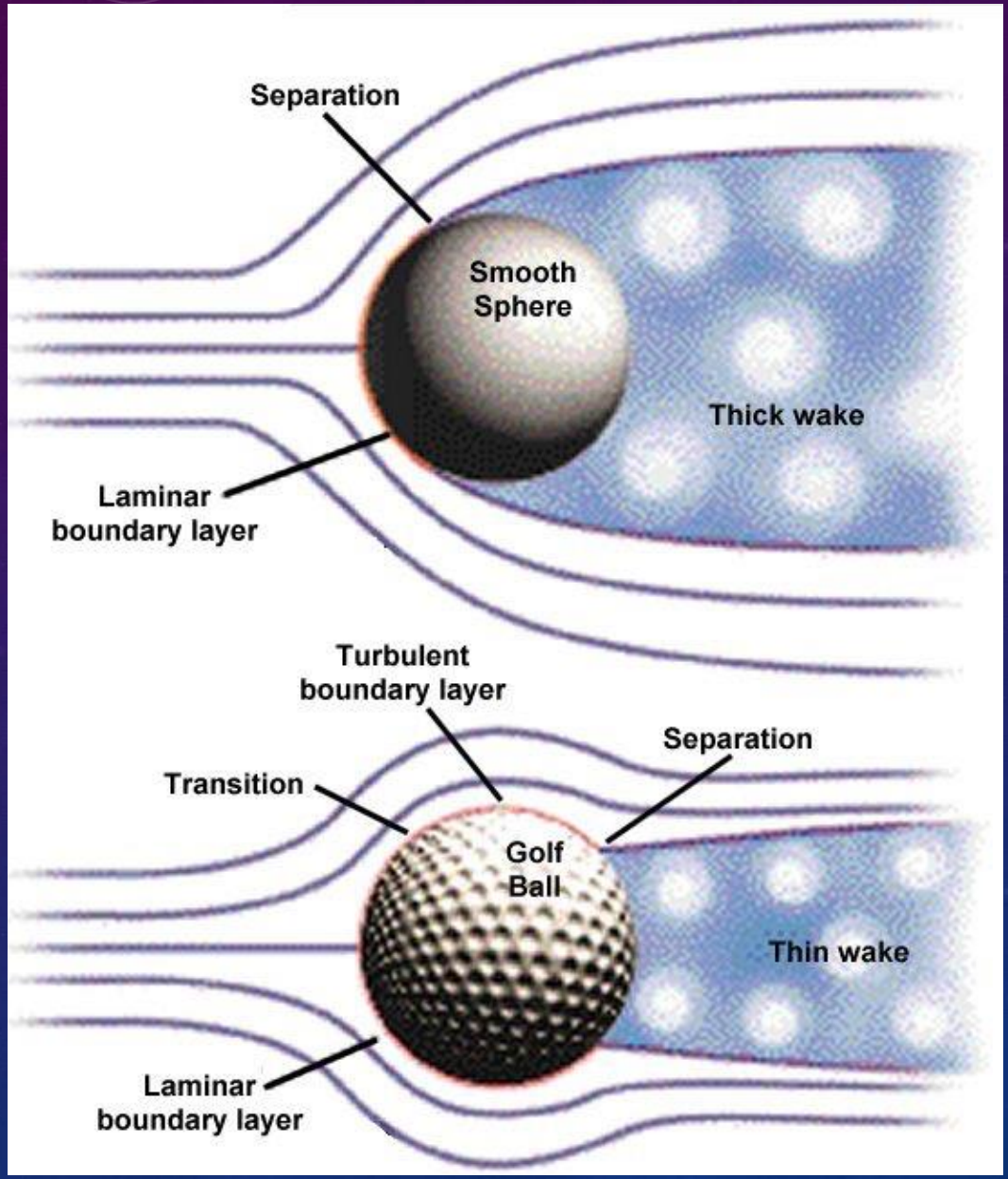

Class 1
Class 2

DATA FLOW


Data
Rand Bell Pareto
Pre-process
Bucket similar things together
Algorithm
CenterPoint or Boundary
Prediction
understand error metric

SINGLE MODEL LIMIT (~TOP 25%)

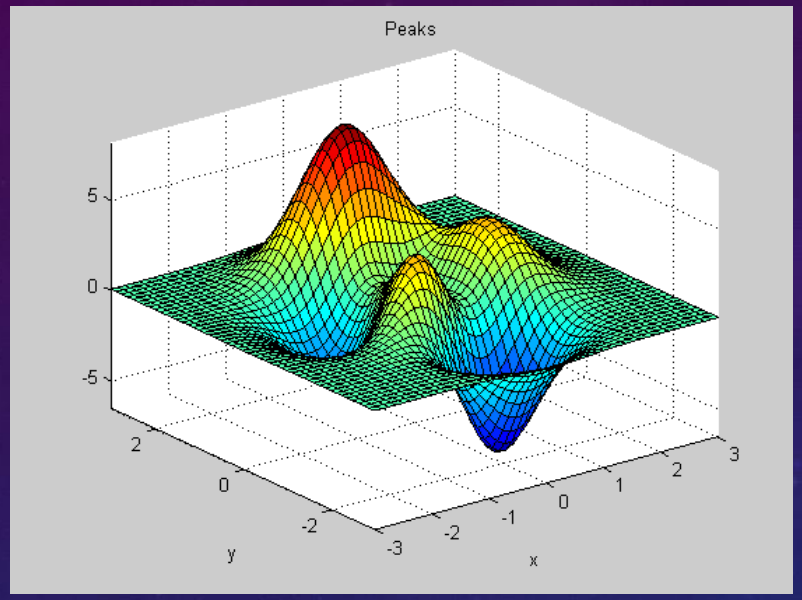


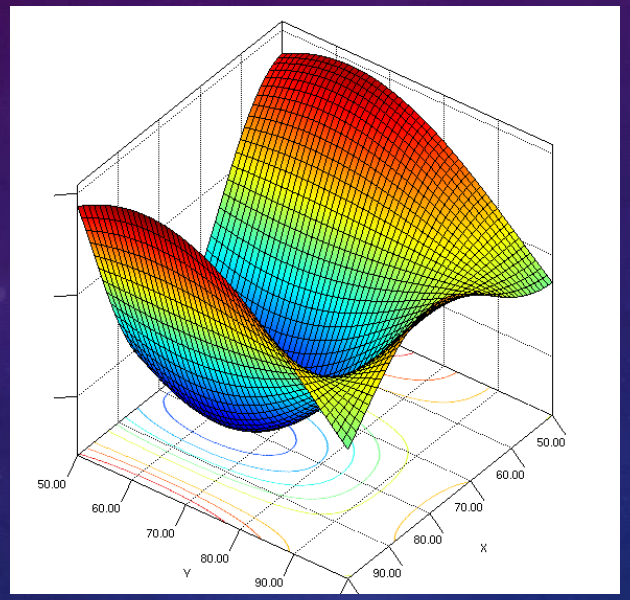

Separation
Smooth Sphere
Thick wake
Laminar boundary layer
Turbulent boundary layer
Separation
Transition
Golf Ball
Thin wake
Laminar boundary layer

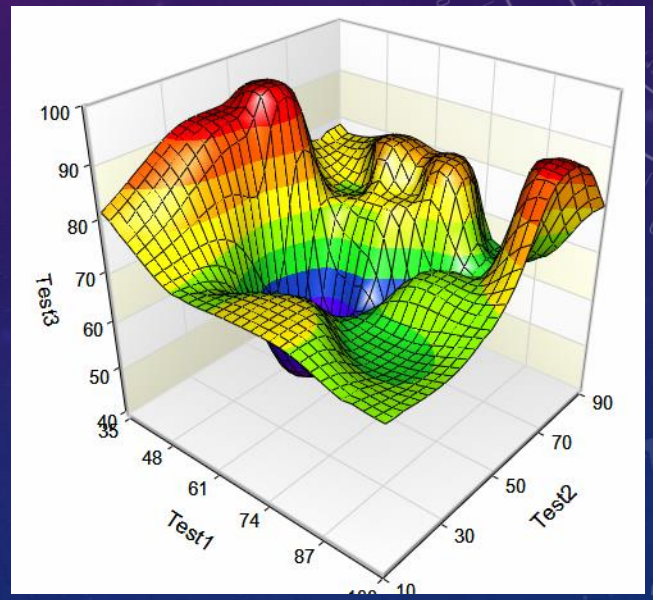
ENSEMBLE / MULTI-MODEL (RANK 25% TO 10%)

- Stable. model weakness complement
- (Imagenet 2016)

MULTI-MODEL (RANDOM SAMPLING/PARALLEL)



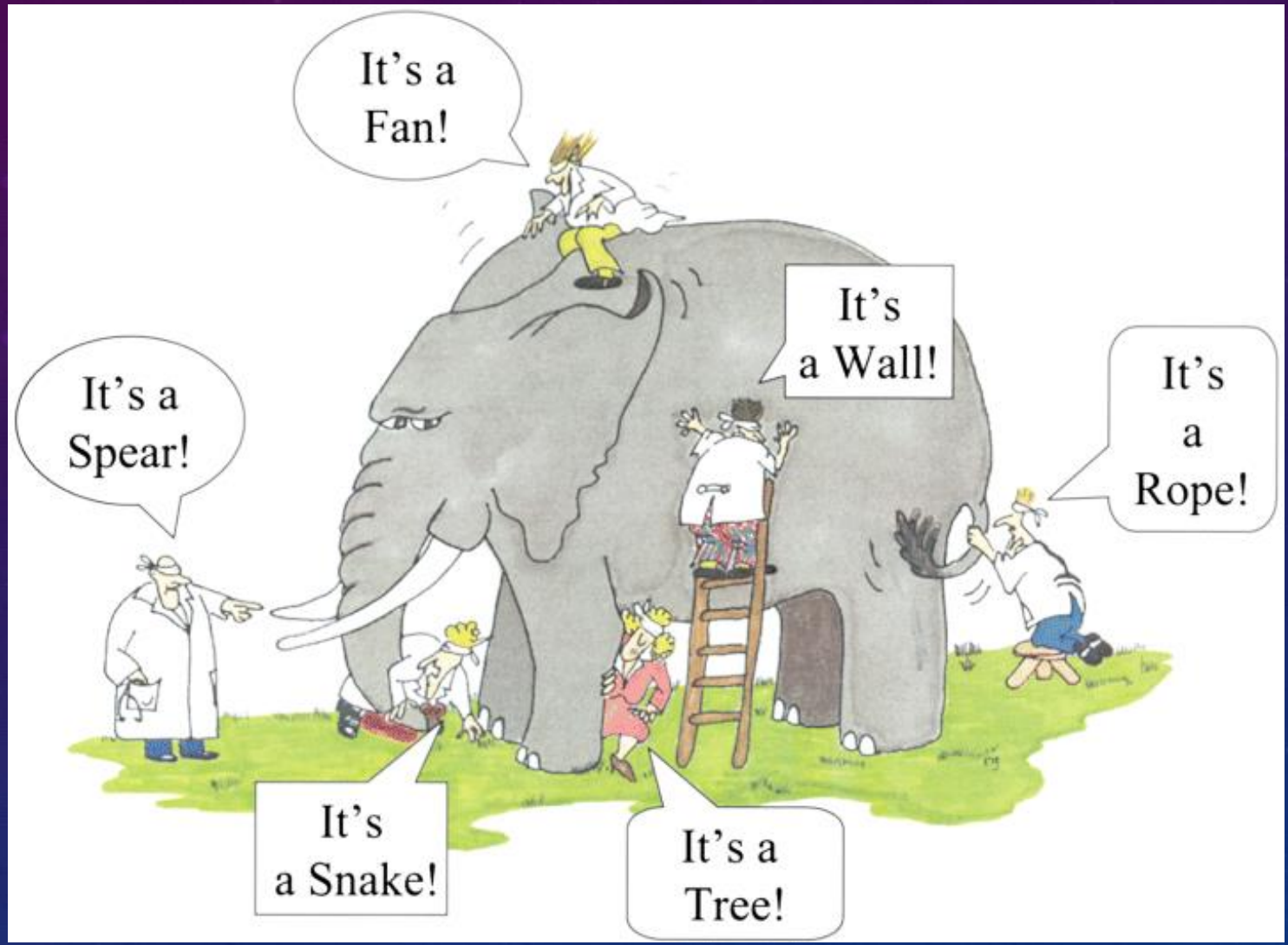




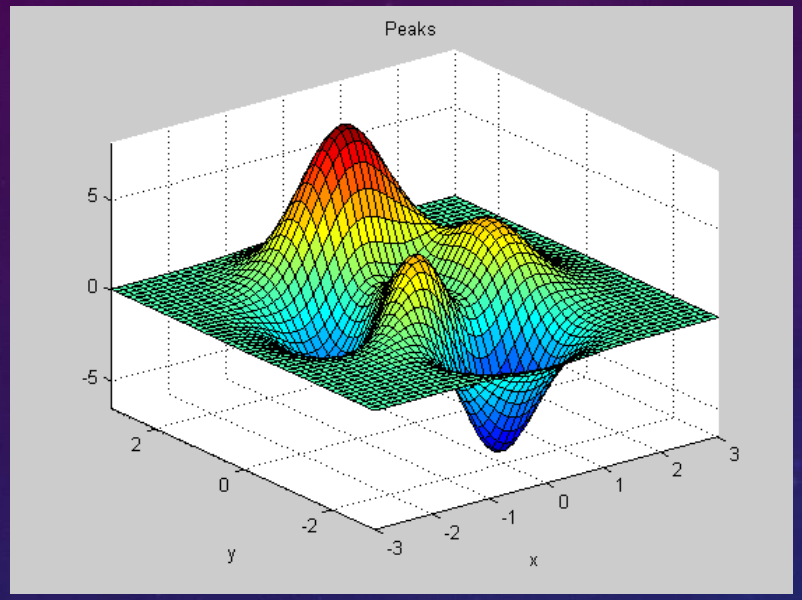
- A1 (dataset 1)

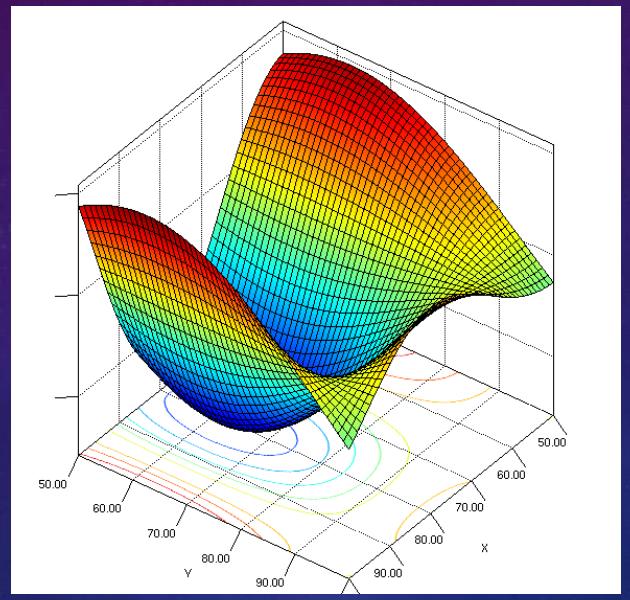
A2 (dataset 2)

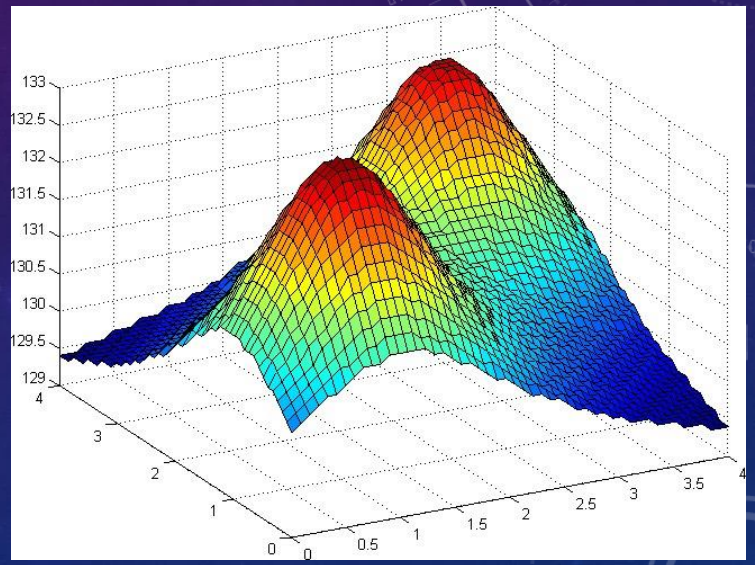
A3 (dataset3)


It's a Fan!
It's a Spear!
It's a Wall!
It's a Rope!
It's a Snake!
It's a Tree!

MULTI-MODEL (BOOSTING/SERIAL)



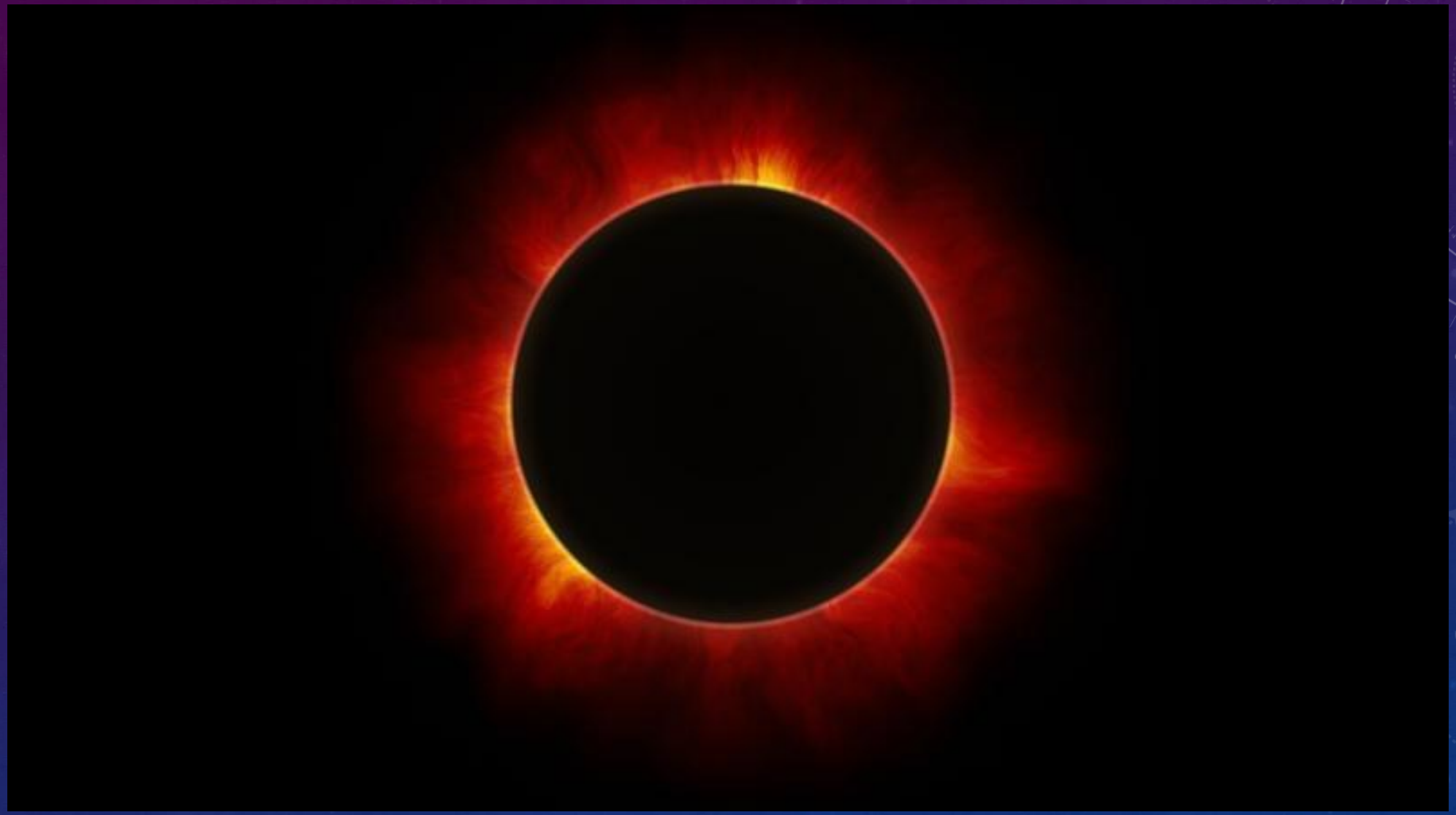




- A1 A2 (Focus on misclassification from A1) A3 (Focus on miss from A1 + A2)

새 학기 **전과목 실력 UP!**

진단평가 대비도 미리 준비하실 수 있습니다!



WORKING IN TEAM (LIKE AN ENSEMBLE)

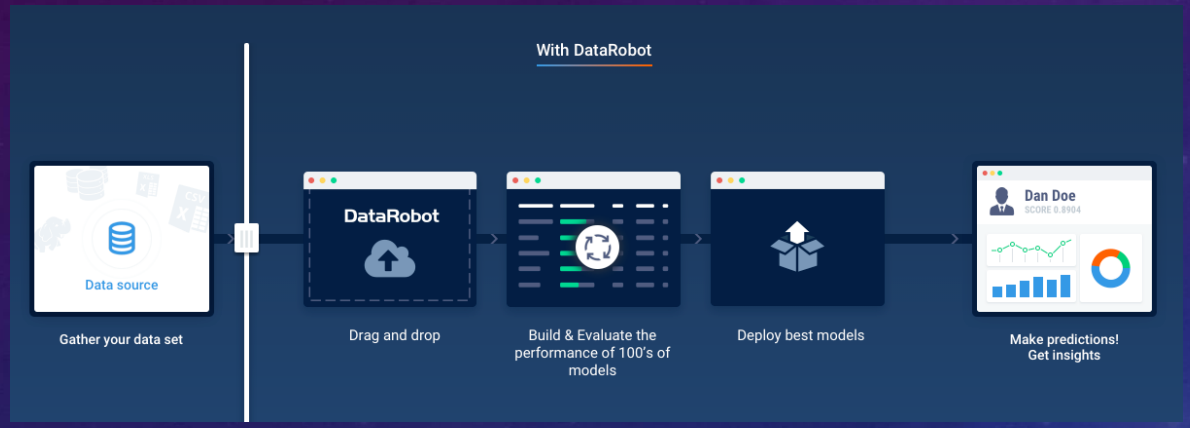
- Merge at the end. Not pipeline.
- Mix different strategy.

TAKING IT EXTREME (TOP 10% TO TOP 1)

TAKING IT TO EXTREME (1)

- Try Every Possible Hyper-Parameter.

AUTOMATED ML / AUTO-COMPETE (끝판왕)



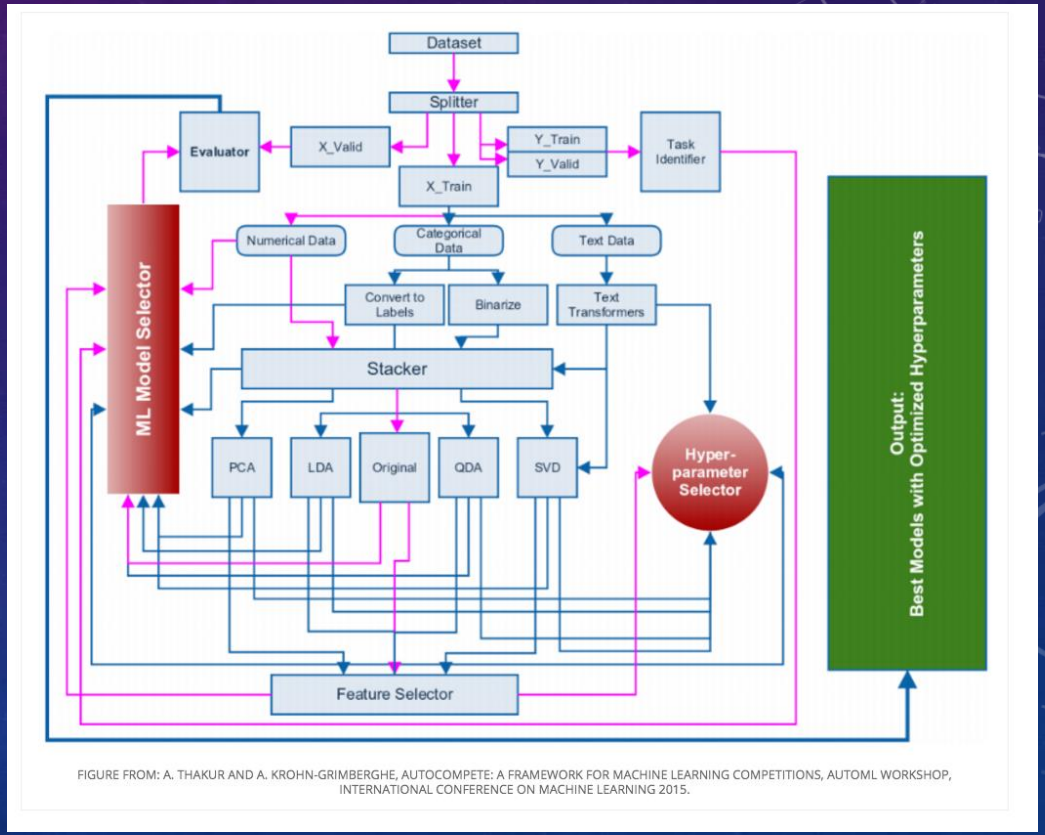
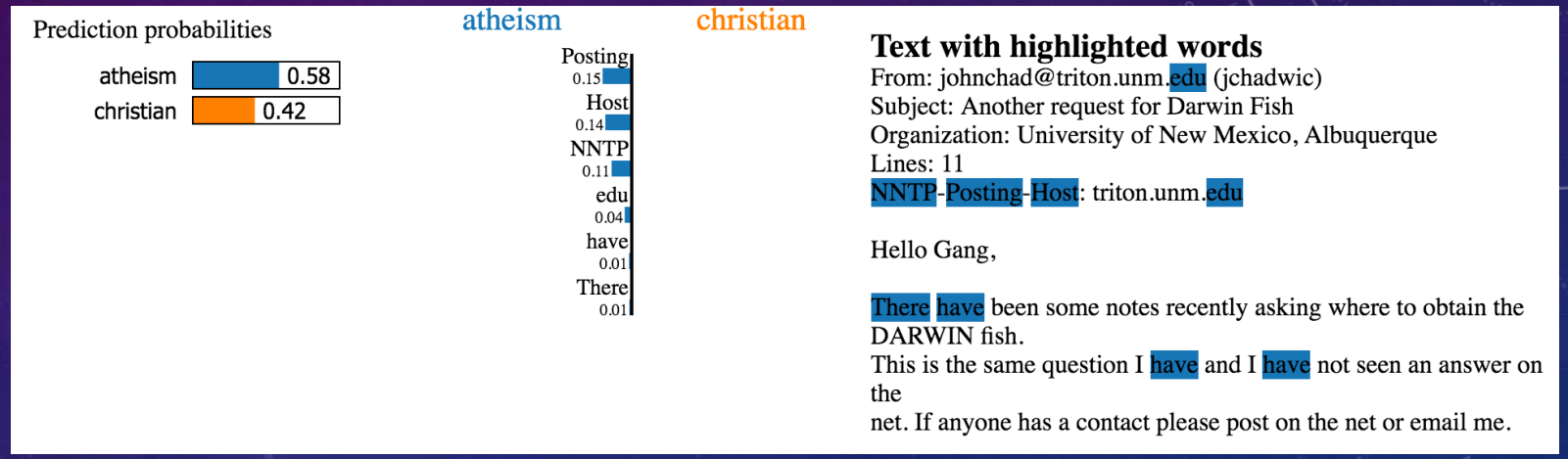


FIGURE FROM: A. THAKUR AND A. KROHN-GRIMBERGHE, AUTOCOMPETE: A FRAMEWORK FOR MACHINE LEARNING COMPETITIONS, AUTOML WORKSHOP, INTERNATIONAL CONFERENCE ON MACHINE LEARNING 2015.

TAKING IT TO EXTREME (2)

- similar performance
- reduced load

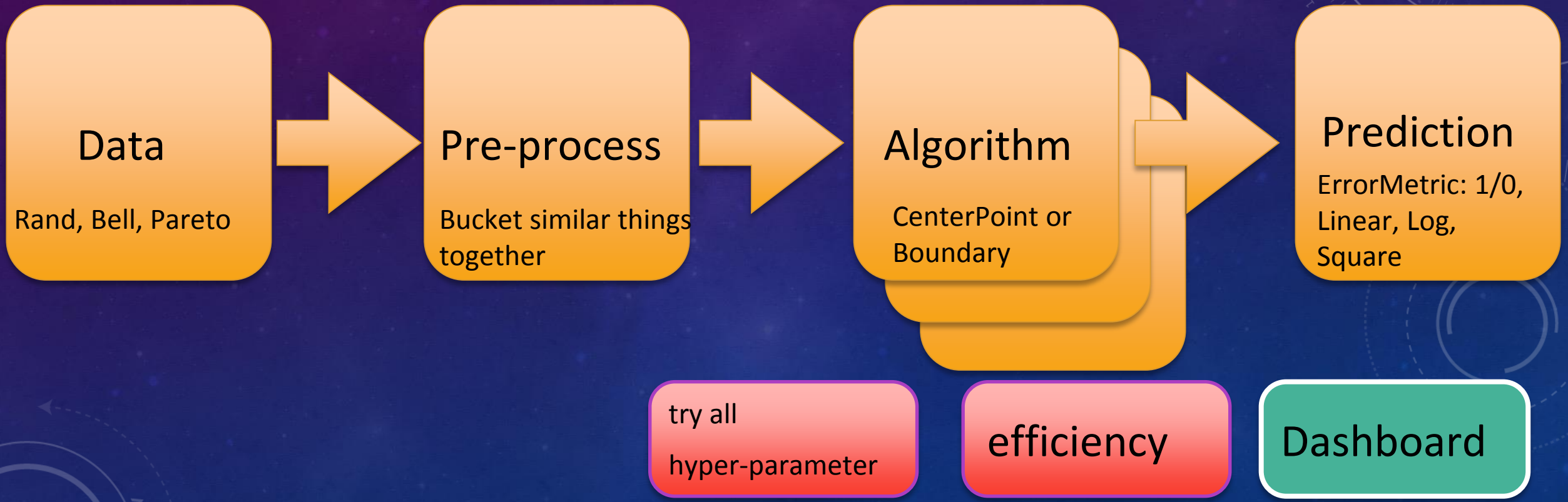
BRING IT BACK DOWN TO EARTH



- Explain it to me like I am a ~~six year old~~ Human.

- To put into a production system with millions of users, you need to convince me.

EVERYTHING PICTURE



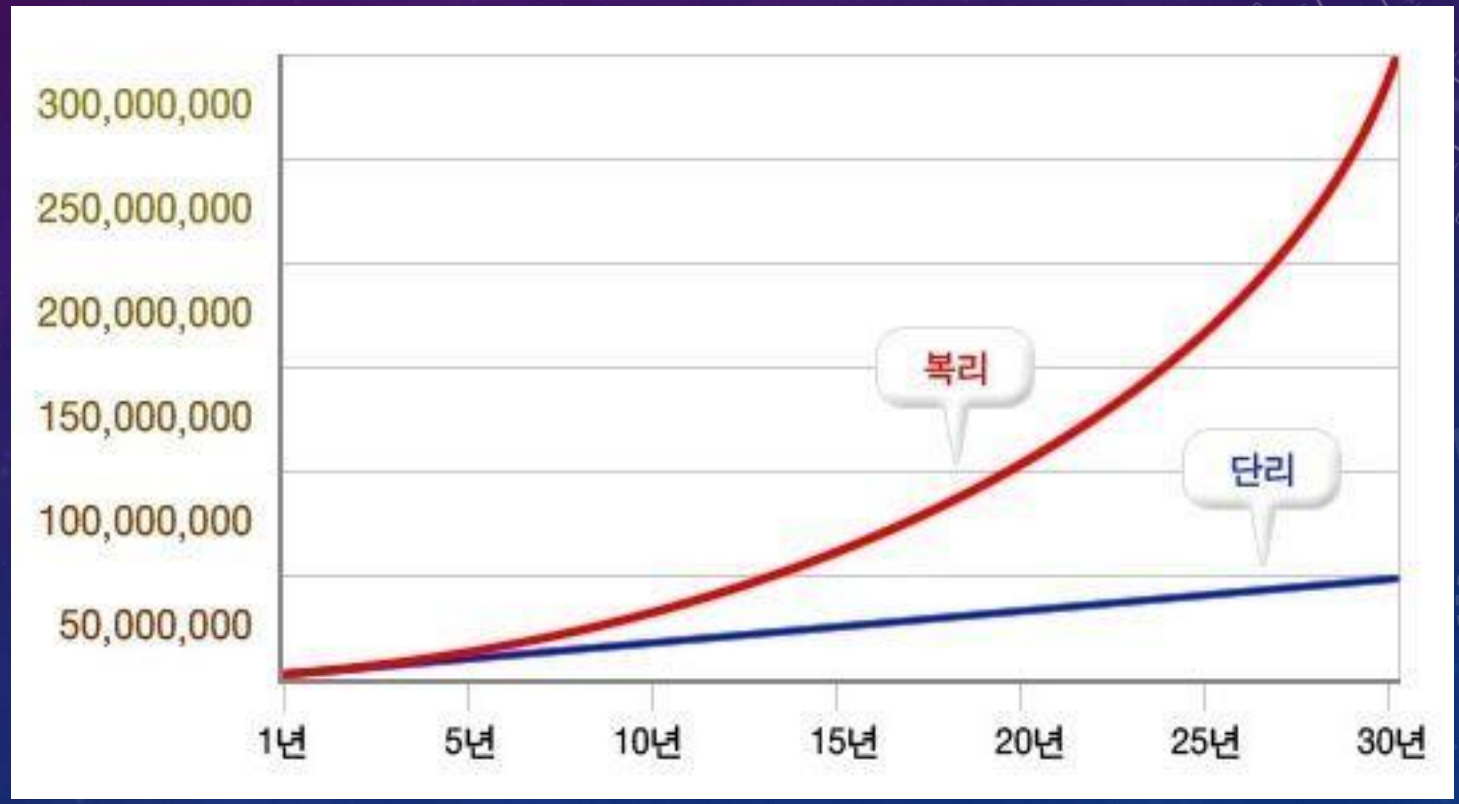
- 제발 어려운거 쓰지말고. 시작은 간단하게.
- 데이터를 빠삭하게.
- 그리고,


THIS IS THE ONLY GAME
THAT IS NOT FAKE.
imgflip.com

진짜 레알 이유는

진짜 레알 이유는

- 지식의 “복리 이자”



THANKS.

- We are hiring.
- Top School OR Kaggle top 25% x2 OR Do “something smart” better than us.