

삼성 오픈 소스 컨퍼런스 2019

청소년도 같이 노는 DATA SCIENCE

선린인터넷고등학교 | 김현우

SAMSUNG OPEN SOURCE CONFERENCE 2019



목차

발표자 소개
About Me

금융 데이터로 놀아보자!
삼성전자 주식 예측 해보기

누구나 쉽게 하는 데이터 사이언스

데이터 사이언스 어떻게 공부해야 할까?

SOSCON 2019
마무리 - Kaggle

SAMSUNG OPEN SOURCE CONFERENCE 2019



About Me



SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

청소년이...
DATA SCIENCE를?

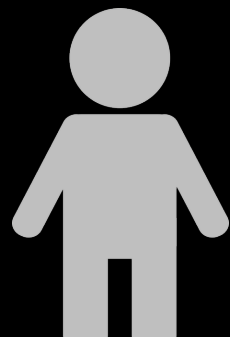
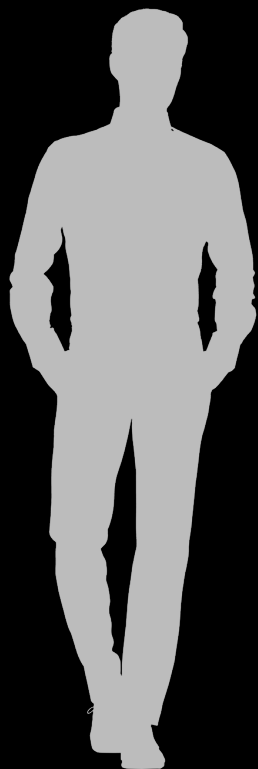


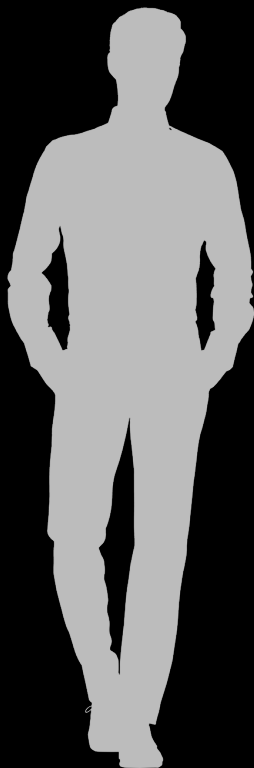
통계학, 선형대수

Machine Learning

Deep Learning

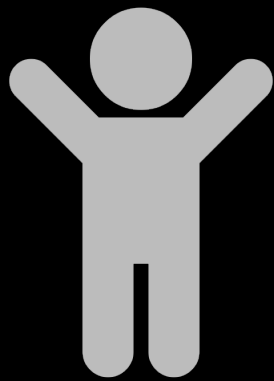








SOSCON2019







삼성전자 주식 예측해보기

금융 데이터로 [▲]놀아보자!

평생 먹고



SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

삼성전자 주식 예측해보기

**‘삼성전자 주식 예측해보기’ 발표는
완성된 프로젝트가 아닙니다.**



SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

삼성전자 주식 예측해보기

현재 진행형 '삼성전자 주식 예측 도전기'

SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019



1 번째 시도



SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

삼성전자 주식 예측해보기

과거의 주가로 LSTM, GRU를 사용해
미래의 주가를 예측할 수 있을까??

SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019



삼성전자 주식 예측해보기

YAHOO!
FINANCE



SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

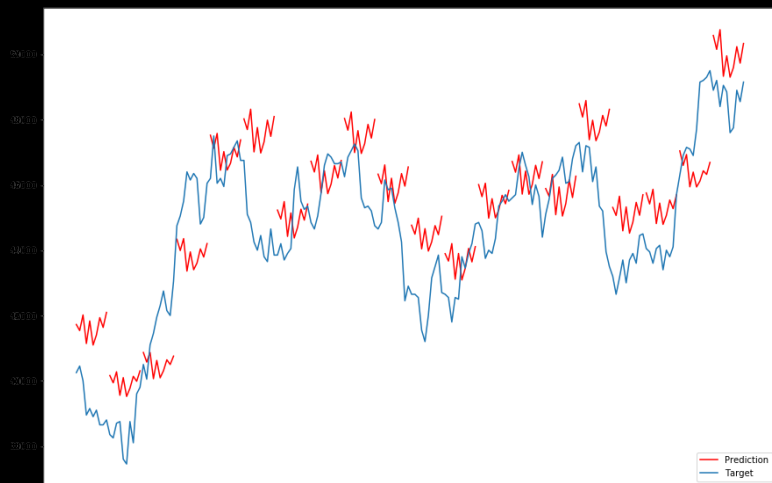
삼성전자 주식 예측해보기



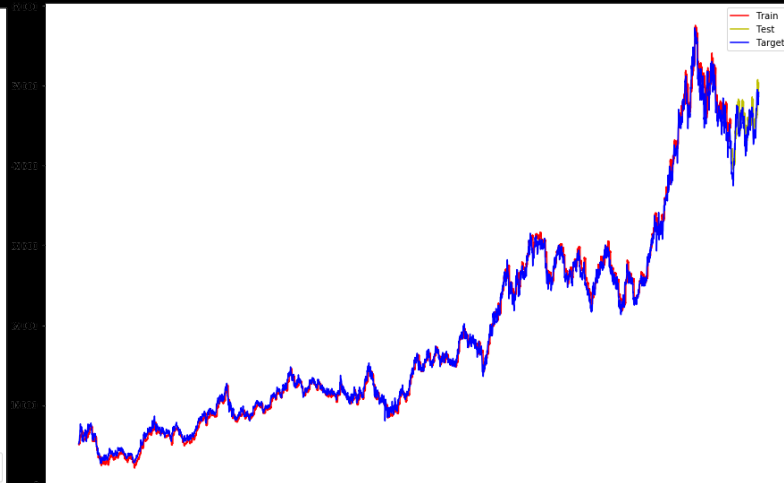
SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

삼성전자 주식 예측해보기



붉은색 : 예측, 파란색 : real data

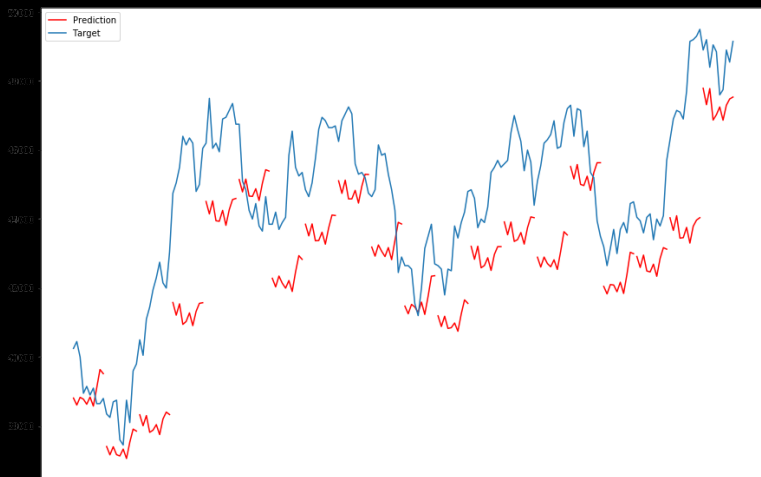


붉은색 : 학습, 노란색 : test
파란색 : real data

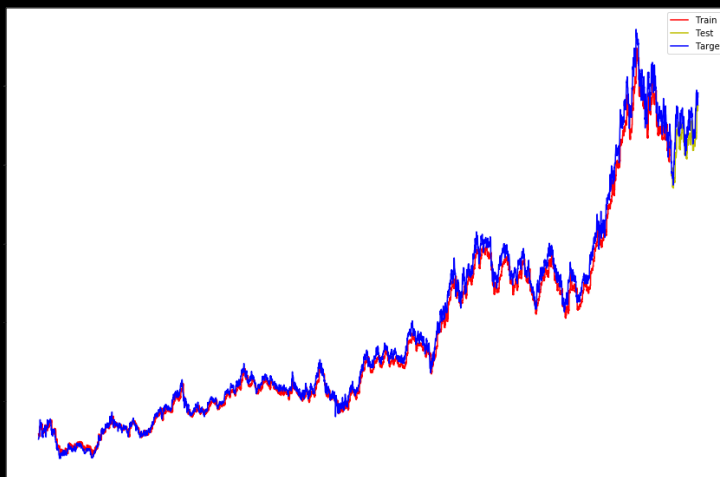
SOSCON2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

삼성전자 주식 예측해보기



붉은색 : 예측, 파란색 : real data

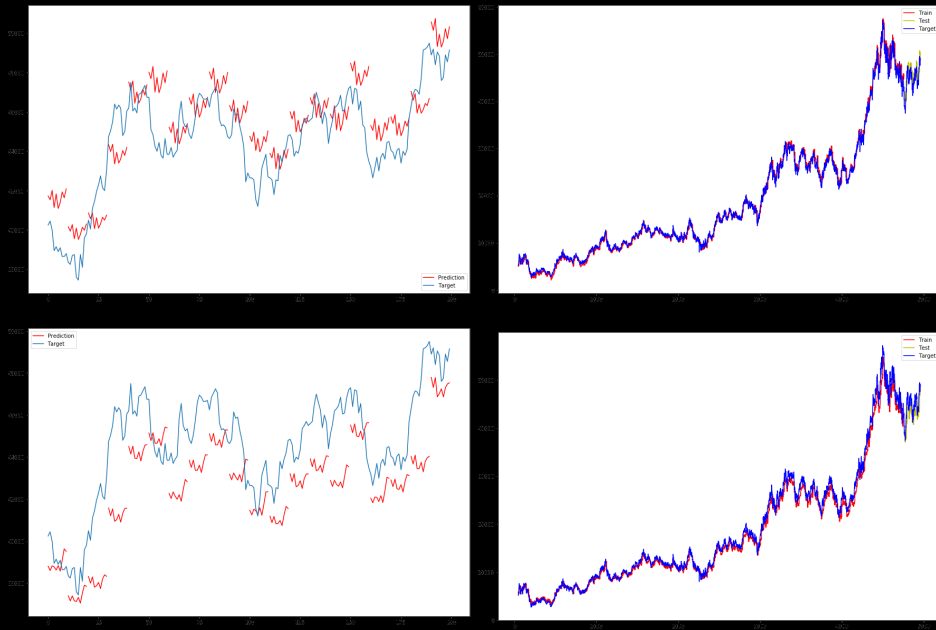


붉은색 : 학습, 노란색: test
파란색 : real data

SOSCON2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

삼성전자 주식 예측해보기



SOSCON2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

삼성전자 주식 예측해보기

?



SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

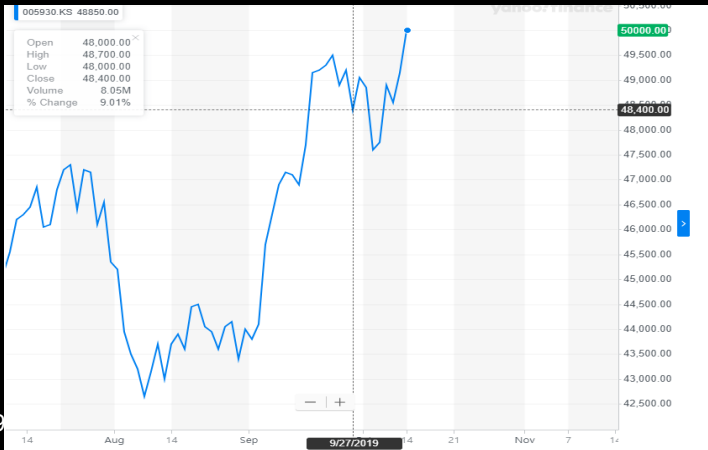
2번째 시도



삼성전자 주식 예측해보기

```
kimhyunwoo ~/Desktop/Programming/Web/samsung $ python3 Web\crawler.py
```

삼성전자, 산골학교부터 작은 도서관까지 '삼성 스마트스쿨' 지원
삼성전자가 산골학교부터 작은 도서관 등 16개 기관에 '삼성 스마트스쿨'을 지원한다고 27일 밝혔다.
'삼성 스마트스쿨'은 정보 접근성이 낮고 디지털 교육이 어려운 환경에 있는 학생들에게 스마트 기기과 솔루션을 지원해 교육 격차를 해소하는 사회공헌 프로그램이다. 매년 공모를 통해 지원 기관을 선발하고 있다.
이날 삼성전자는 따르면, 올해는 지난 4월부터 공모를 시작한 400여개 기관이 응모했고, 3차례에 걸친 심사과 온라인 투표를 통해 최종 16개 기관이 뽑혔다. 이 중에는 지리산 산골마을 분교인 '금서초등학교', 탈북청소년 대안학교인 '우리들학교', 강원도 영월군의 폐광지역에 위치한 '별마로 작은 도서관' 등이 포함됐다.
삼성전자는 지난 26일 서울 '느헤미야코리아 다움학교', 강원도 '별마로 작은 도서관', 전라남도 '천태초등학교' 등 3개 기관에서 '스마트 스쿨' 개소했고, 다음달까지 16개 기관 모두 운영에 들어갈 예정이다.
삼성전자는 2012년부터 '삼성 스마트스쿨'을 도서벽지 학교에 지원했다. 2016년부터는 지역에 상관없이 학교뿐만 아니라 병원, 다문화가족지원센터, 지역아동센터, 특수학교 등의 모든 교육 기관으로 지원을 확대했다. 지난 해까지 전국 83개 기관에 스마트스쿨을 지원했다.
삼성전자는 디지털 플립차트 '삼성 플립', 태블릿, 노트북 등 스마트 기기를 지원하고, 환경에 따라 적합한 교육 콘텐츠도 제공한다. 아울러 삼성전자 임직원들이 직접 멘토로 참여해 지원 기관별 맞춤형 교육 프로그램을 기획하고, 스마트스쿨이 안정적으로 운영되도록 교육 봉사활동도 실시한다.
지난해에 이어 올해 두번째 스마트스쿨 멘토로 참여한 삼성전자 무선사업부 조성진 엔지니어는 "어렸을 때 좋은 멘토를 만나 어려운 시기를 잘 넘겼던 경험이 있다"며, "지역적, 경제적 교육 격차가 심해지는 상황에서 스마트스쿨을 통해 교육 격차를 줄이는데 보탬이 되도록 최선을 다하겠다"고 말했다.
한편, 삼성전자는 지난 2월 새로운 사회공헌 비전인 '함께가요 미래로! Enabling People'을 발표하고, 청소년 교육을 주요 사회공헌 테마로 임직원들이 참여하는 다양한 활동을 펼치고 있다.
검찰 내부 사정에 밝은 전현직 관계자들에 따르면 삼성바이오로직스 분식회계 의혹을 수사 중인 서울중앙지검 특수4부(부장 이복현)는 최근 들어 '물음표 공시 누락'과 '삼성 합병 이슈'를 다시 살펴보고 있는 것으로 알려졌다.
검찰은 지난해 12월13일 대규모 압수수색을 기점으로 상반 분식회계 의혹 수사에 나섰다. 올해 2월에는 서울중앙지검 특수2부를 전담 수사팀...
소상공인방송이 소상공인들의 역경과 도전의 비하인드 스토리를 담은 '소공인 작은 거인들' 종합편을 이달 24일 밤 9시부터 방송할 예정이라고 14일 밝혔다. 소상공인방송정보원이 운영하는 '소상공인방송'은 국내 유일의 소상공인과 전통시장 분야 전문 방송채널로서 지역경제 활성화와 전통시장의 부흥을 도모하는 다양한 제작 프로그램을 선보이고 있다. '소공인 작은 거인들...'
불경기에서 벗어나길 줄 모르고 서민들 등골만 휘는 이 때, 전통시장에서 대박을 친 아이미들이 있다. 연세시장(서울 은평구)에 가면 오직 한 가게만 긴 줄이 늘어서 있다. 이 곳은 다른 아닌 집에 가서 바로 조리해서 먹을 수 있게 만든 즉석 찜개 포장 전문점 '보글보글'해가죽화 시대에 맞춰 깔끔하게 소리를 건곤교장 하는 것이 이 집만의 차별화 된 비법이다. 직접 ...
서울상공회에서 10월 12일과 13일 '작품명:서울술'을 주제로 전시·공연 행사가 진행된다. 이번 행사는 올해로 14번째를 맞는 서울술 가을페스티벌의 일환이다. 올해 서울술 가을페스...



SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019



네이버 기사의 긍/부정 판별

: 큰 사건이 아닌 이상 기사의 긍/부정과 주가의 상/하락 상관관계를 파악하기 힘들다



삼성전자 주식 예측해보기

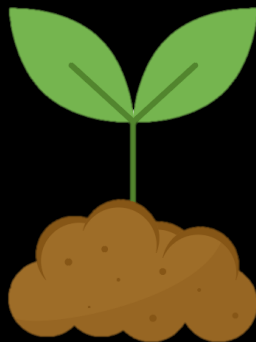
FAIL



SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

삼성전자 주식 예측해보기



SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

삼성전자 주식 예측해보기

틱 데이터 활용

재무제표 데이터 활용

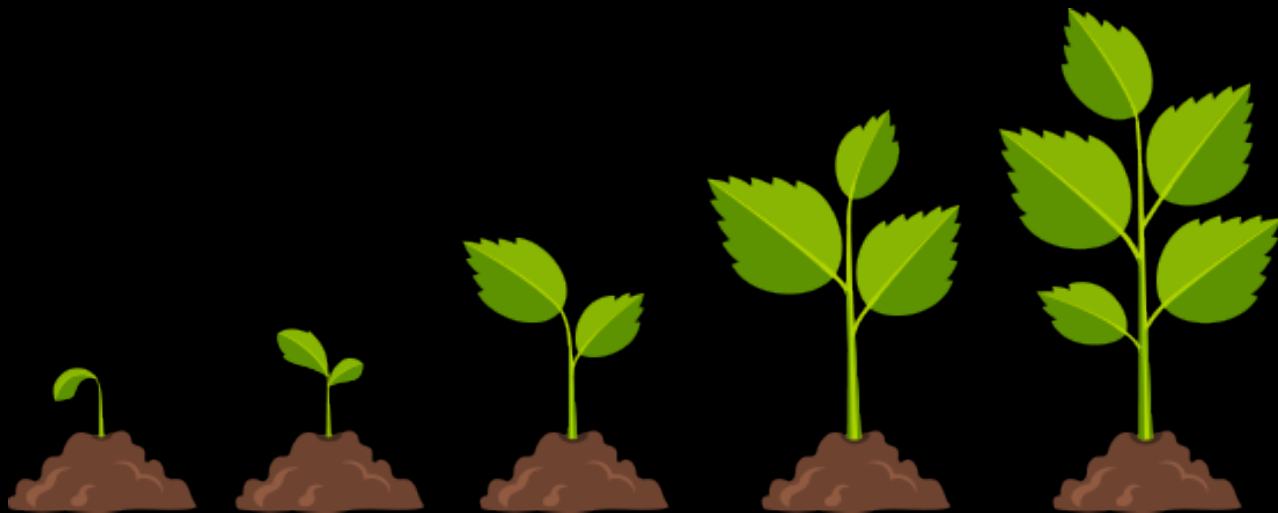
딥러닝 알고리즘 시도

SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019



삼성전자 주식 예측해보기



SOSCON2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

삼성전자 주식 예측해보기



SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

삼성전자 주식 예측해보기

CHATBOT PROJECT



파이썬 크롤러 모듈



엔진 선택



...

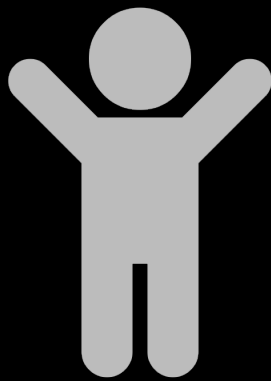
SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019



삼성전자 주식 예측해보기

DATA SCIENCE



SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

Data Science 어떻게 시작할까?

누구나 쉽게 할 수 있는 Data Science



SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

직접적인 내용 X

무엇을 어떻게 공부해야 하는가?



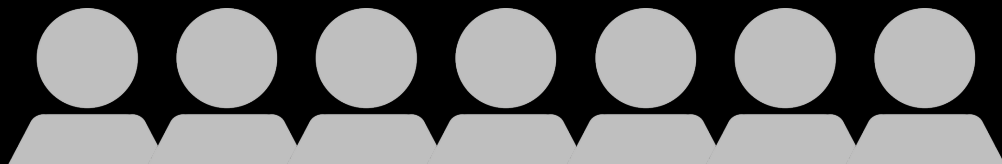


How

어떻게 공부하나요?

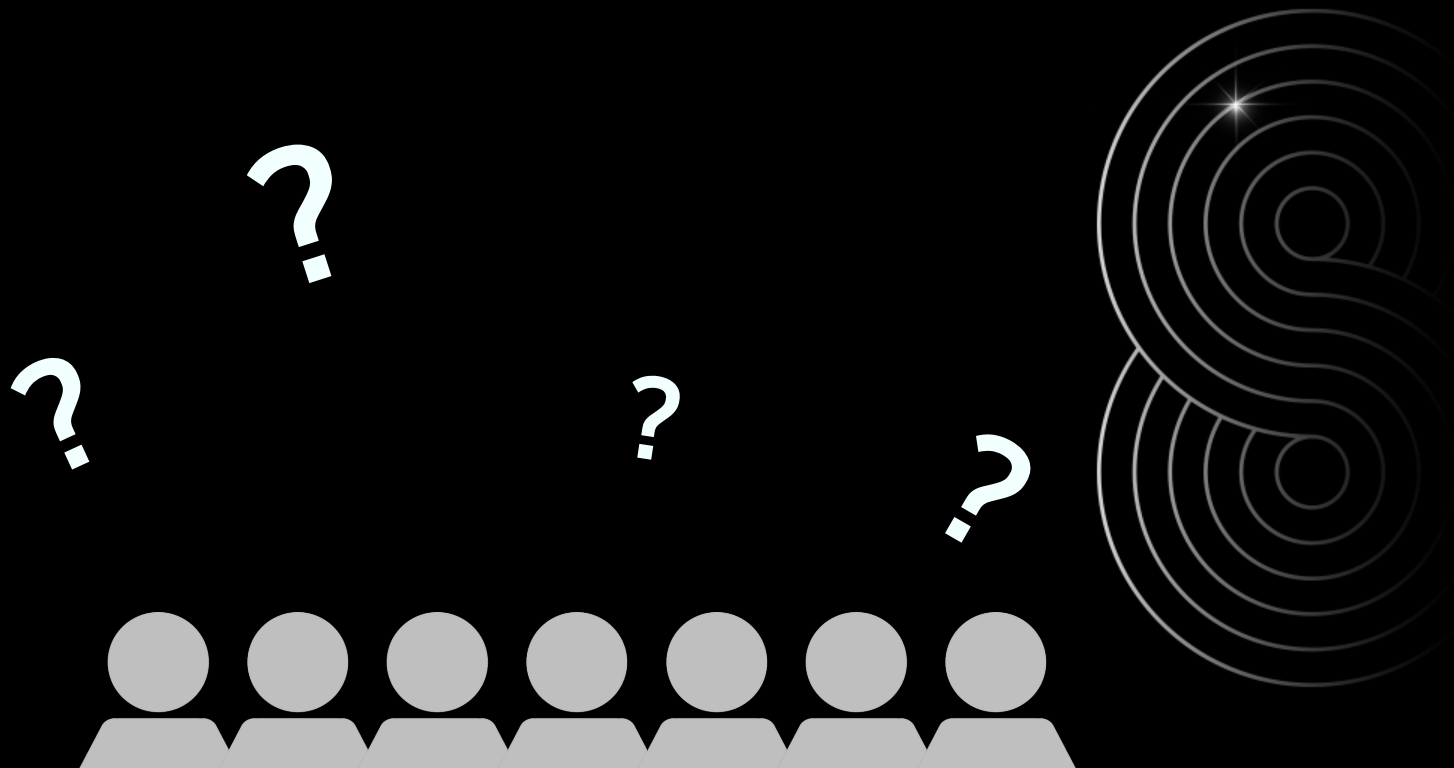


根本



Why
왜 공부하나요?





My Story

나의 이야기





당시의 나, 17세







Google



선배



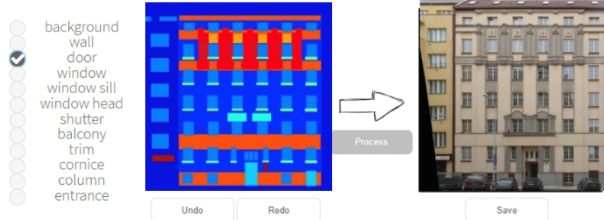


Image to Image

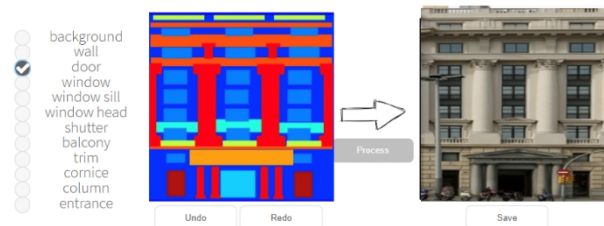
Pix2Pix_Facade

Made by 고동현우[Emotion]

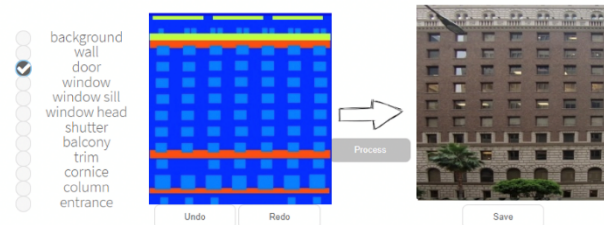
여기다 그림그리면 되는거임~

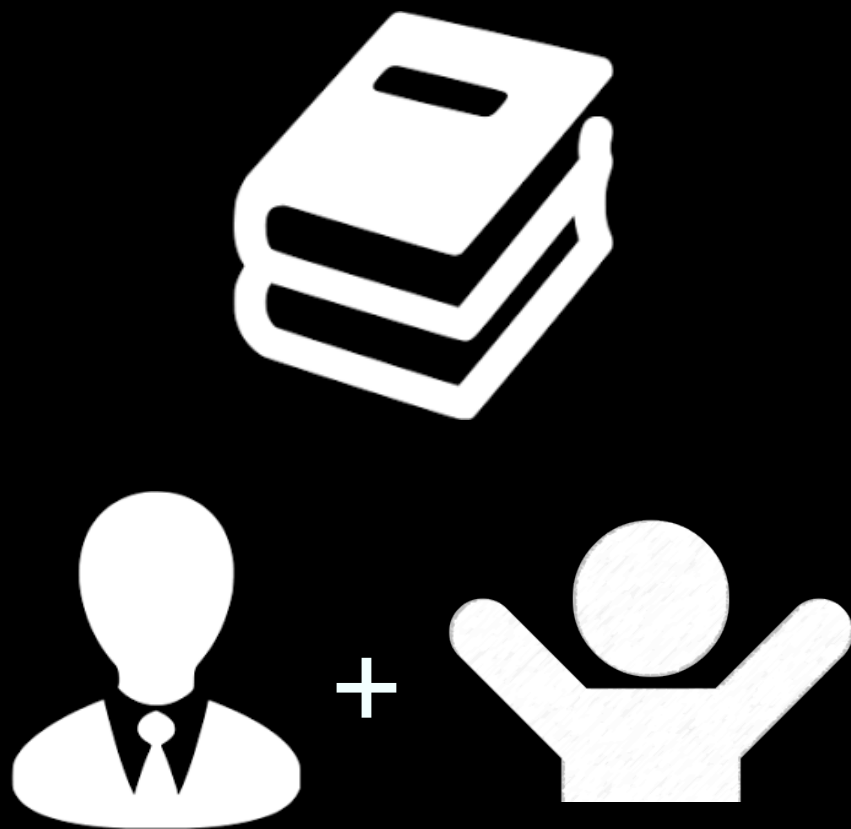


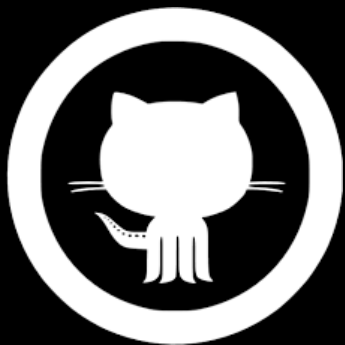
여기다 그림그리면 되는거임~



여기다 그림그리면 되는거임~



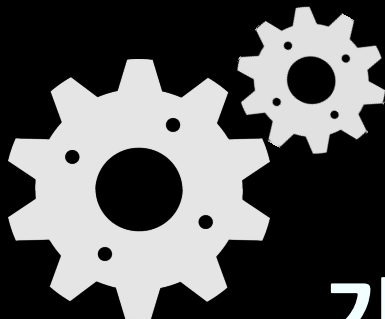




DS / ML / DL



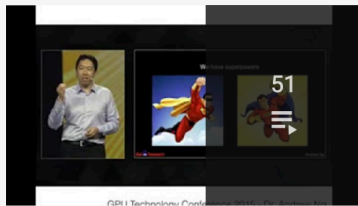
학문



기술







모두를 위한 딥러닝 강좌 시즌 1

Sung Kim

Lec 00 - Machine/Deep learning 수업의 개요와 일정 • 10:05

ML lec 01 - 기본적인 Machine Learning 의 용어와 개념 설명 • 12:29

모든 재생목록 보기

Sung Kim님 팬이메용

Google

cnn 알고리즘

전체

이미지

동영상

뉴스

쇼핑

더보기

설정

도구

검색결과 약 669,000개 (0.47초)

[번역] 딥러닝 (CNN) 직관적 이해 - (1) - HAMA 블로그 - Tistory

<https://hamait.tistory.com> > ...

2016. 4. 9. - 평소 무엇인가를 쉽게 설명하는 능력이 있다고 생각해서, CNN (convolutional neural network) 도 그렇게 해볼까 했는데 ... 오류역전파 알고리즘.

컨벌루션 신경망 (Convolutional Neural Networks, CNN) ~ 개요

<https://horae.tistory.com> > entry > 컨벌루션-신경망-Convolutional-Neural... >

2017. 3. 6. - 영상과 음성에서 좋은 성능을 보이는 알고리즘으로 CNN이 있다. 이 합성곱신경망(컨벌루션 신경망, CNN)은 전처리를 추가한 다층퍼셉트론의 한종 ...

[머신 러닝/딥 러닝] 합성곱 신경망 (Convolutional Neural ...

<https://untitledblog.tistory.com> > ... >[머신 러닝/딥 러닝] 합성곱 신경망 (Convolutional Neural Network, CNN)과 학습 알고리즘. Bool
2019.02.06 17:00. 1. 이미지 처리와 필터링 기법. 필터링은 이미지 ...

CNN, Convolutional Neural Network 요약 - taewan.kim 블로그

taewan.kim > post > cnn >

2018. 1. 4. - 이미지의 공간 정보를 유지한 상태로 학습이 가능한 모델이 바로 CNN(Convolutional Neural Network)입니다. CNN(Convolutional Neural ...

CNN의 주요 용어 정리 · CNN 구성 · CNN 입출력, 파라미터 계산

컨벌루션 뉴럴 네트워크란? - MATLAB & Simulink - MathWorks

<https://kr.mathworks.com> > deep-learning > convolutional-neural-network >

컨벌루션 뉴럴 네트워크(CNN 또는 ConvNet)는 딥러닝에 가장 많이 사용되는 알고리즘 중 하나입니다. MATLAB 예제와 도구를 사용하여 컨벌루션 뉴럴 네트워크를 ...

딥러닝 - 초보자를 위한 컨벌루션 뉴럴 네트워크를 이용한 이미지 ...

<https://bcho.tistory.com> > ... >

2016. 11. 30. - CNN은 전통적인 뉴럴 네트워크 앞에 여러 계층의 컨벌루션 계층을 ... 앞에서 설명할 때, softmax 나 뉴런에 대한 세부 알고리즘 ReLu 등과 ...



Multi Layered Neural Network

수용 영역(Receptive field)

566 LeCun, Rosen, Olivier, Houdouin, Houdouin, and Jodanis



이 자료는 Gradient-Based Learning Applied to Document Recognition에 소개된 LeNet-5의 구조입니다.

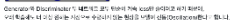
위 사진을 보아도 input image에서 Convolutional 연산과 Subsampling을 반복해 주다가, 항상 64번의 크기로 줄어든다. 이후 Fully-connected 네트워크이다. 이러한 구조에서 CNN의 특징이언제든, 다른 곳에 많이 있음, 학습됨, 복원된 Convolutional Layer과 Pooling Layer(Subsampling)를 중심으로 Layer에서 적용되는 연산과 적용되는 연산의 방법을 정리해보도록 하겠다.

Convolutional Layer는 Convolution 연산을 거치는 층으로써, Convolution 연산에 필요한 용어들(필터, 커널, 슬라이딩 윈도우, 스트라이프)을 알아야 한다.

생성 분포: Data generating distribution, **판별 분포**: Discriminator distribution
배치 선: Generative distribution, **인코딩 변환**: $x \rightarrow G(z)$ mapping
Latent, **Lat**, **z**, **잠재**는 학습을 통해 인코더에서 추출 가능한 수식 값이다.
 실제 생성된 데이터와 학습 데이터를 비교할 때의 불일치를 의미하며, 가중치에 따라
 판별 손실(Generator loss)과 판별 정확도(Disc) = 옳고(부름) 틀림(가)를 나타낸다.

생성적으로 학습이 이루어졌을 때의 모습입니다.

모양의 진동(Oscillation)



Mode Collapsing의 실행을 돕기 위해 MINST도 해당 용어에 대해 설명한다.

© 2006 Blackwell Publishing Ltd, *Journal of Internal Medicine* 260: 409–416

이런 GAN의 특징을 이해하기 쉽고 재밌어서 많은 사람들이 GAN을 열연한 데 쓰는 비유도 자리매김했습니다.

원어 위조(조작)범은 신해 사제보 목공비 남용에서 경한이 구별할 수 있게 하는 것이 목공비이고, 경한은 뒤조작(가짜)한 뒤조작 자와 실과 자를 붙인 것이 구별하는 것이 다르다. 이의 구별하면 경한의 구별하는 경한으로 뒤조작 범을 화수화, 경한을 화수화 하는 것. 이 국문가 김(10)



Mode



예를 들어는 구조야고 부르는 것일까요? 왜 하면 *Domini sunt Genesim*가 아닌 *Genesim sunt Domini*라는 식이 붙어가지고, *Genesim*은 *Genesim*을 *Recti*라고 부르는 것이 아니라 *Domini*가 부르는 것이 맞는지 모르겠습니다.

© 2014 Generative Adversarial Nets (2014) by the author. All rights reserved.

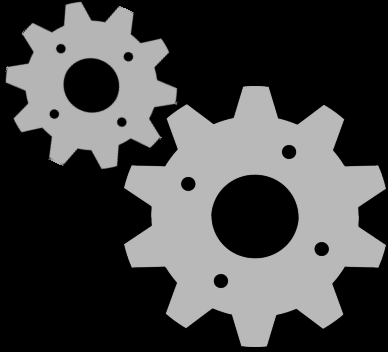
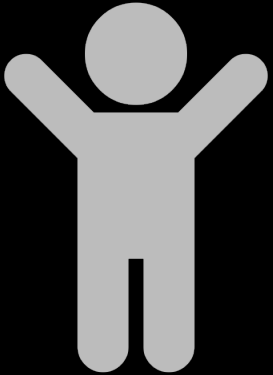
연구 결과 많은 학파적 관습과 이론을 뒤집기 위해 진념으로 헌신하고 있음에도



SOSCON 2019

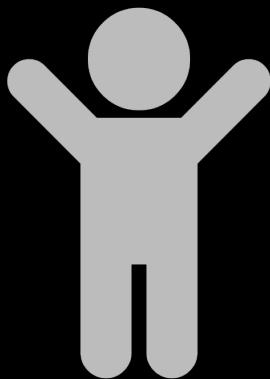


github.com
/hwk06023



이해도





book

SOSCON2019



물론 이것들을 아직 다 보진 않았습니다. 😊

 **fast campus**

정가

349,000 원

할인 금액

265,000 원

결제 금액

84,000 원



1. 본인이 공부하고자 하는 목표를 명확히 하자.
2. 기술을 단순히 사용해 보며 공부하는 것으로 그치지 말고, 이를 이해하고 사용할 수 있도록 기초 원리부터 차근차근 공부하자.



이렇게 열심히 공부했으면..!

대회 해커톤 공모전





미루지 말자



Kaggle

모두가 직접 지금 바로 시작하는 데이터 사이언스



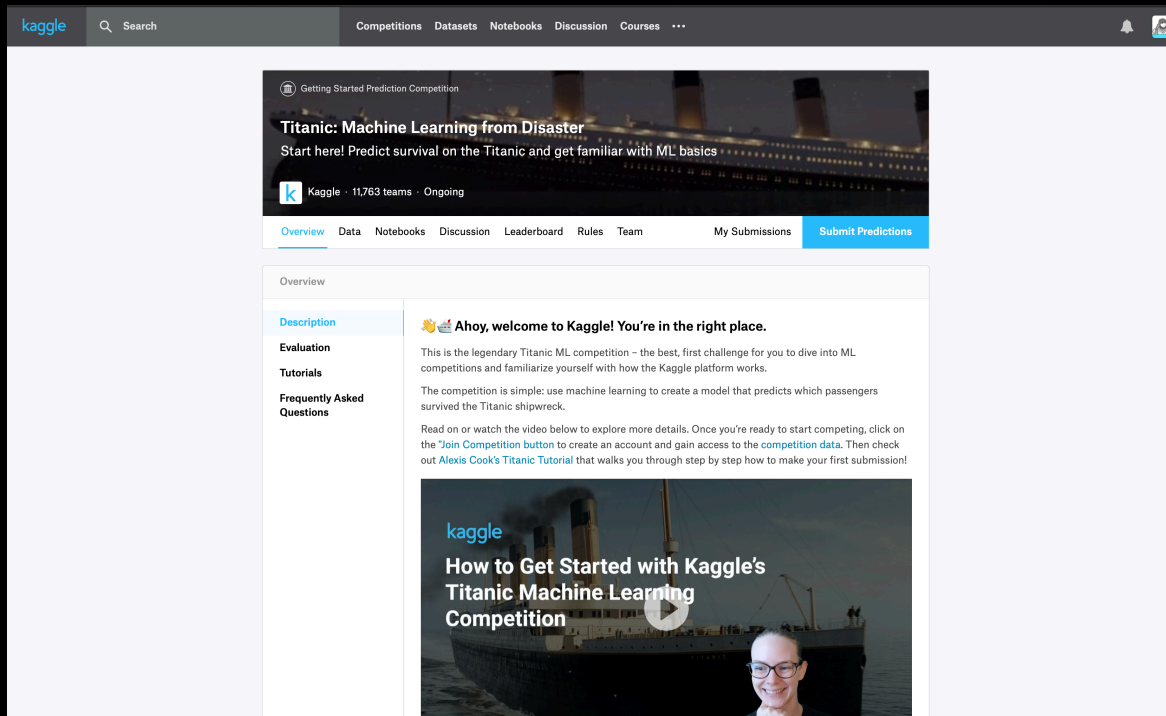
SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019





The screenshot displays the Kaggle website interface. At the top, the navigation bar includes the Kaggle logo, a search bar, and links for Competitions, Datasets, Notebooks, Discussion, and Courses. The main content area features a banner for the 'Titanic: Machine Learning from Disaster' competition, which includes the title, a brief description, and the number of teams (11,763) and the status (Ongoing). Below the banner, there is a sidebar with links for Overview, Data, Notebooks, Discussion, Leaderboard, Rules, Team, My Submissions, and Submit Predictions. The main content area shows a welcome message from Kaggle, a description of the competition, and a video titled 'How to Get Started with Kaggle's Titanic Machine Learning Competition'.

Titanic: Machine Learning from Disaster
 Start here! Predict survival on the Titanic and get familiar with ML basics

Kaggle · 11,763 teams · Ongoing

[Overview](#) [Data](#) [Notebooks](#) [Discussion](#) [Leaderboard](#) [Rules](#) [Team](#) [My Submissions](#) [Submit Predictions](#)

Overview

Description

Evaluation

Tutorials

Frequently Asked Questions

Ahoy, welcome to Kaggle! You're in the right place.

This is the legendary Titanic ML competition – the best, first challenge for you to dive into ML competitions and familiarize yourself with how the Kaggle platform works.

The competition is simple: use machine learning to create a model that predicts which passengers survived the Titanic shipwreck.

Read on or watch the video below to explore more details. Once you're ready to start competing, click on the [Join Competition button](#) to create an account and gain access to the [competition data](#). Then check out [Alexis Cook's Titanic Tutorial](#) that walks you through step by step how to make your first submission!

How to Get Started with Kaggle's Titanic Machine Learning Competition

Google kaggle titanic

검색결과 약 148,000개 (0.35초)

Titanic: Machine Learning from Disaster | Kaggle
<https://www.kaggle.com/titanic> > titanic > 이 페이지 번역하기
 Ahoy, welcome to Kaggle! You're in the right place. This is the legendary Titanic ML competition - the best, first challenge for you to dive into ML competitions ...
 Titanic: Machine Learning from ... > Data > Titanic Data Science Solutions > Kernels

Titanic: Machine Learning from Disaster | Kaggle
<https://www.kaggle.com/titanic/overview> > overview > 이 페이지 번역하기
 Titanic: Machine Learning from Disaster. Start here! Predict survival on the Titanic and get familiar with ML basics. Kaggle: 11,423 teams. Ongoing.

Kaggle과 Titanic research로 시작하는 데이터 분석 - Ascent ...
<https://developers.ascentnet.co.jp/2017/11/24/kaggle-process-review> >
 2017. 11. 24. - 데이터 분석 플랫폼인 Kaggle에 대한 소개와 데이터 분석 기초에 많이 사용되는 'Titanic: Machine Learning from Disaster' 자료를 분석하는 과정에 ...

Kaggle 타이타닉(Titanic) 80.3% 성공 후기 및 코드
<https://doorw.tistory.com/> > ... >
 2018. 6. 4. - 언뜻하셔도 문법우입니다. 최근 데이터분석, 인공지능 분야에 관심이 있어서 스타터를 시작하여 kaggle문제를 풀어보기 시작했습니다. 개인적으로 ...

타이타닉 튜토리얼 1 - Exploratory data analysis ... - Kaggle-KR
<https://kaggle-kr.tistory.com/> > ... >
 2018. 6. 28. - 만약 데이터 사이언스 마스터님 또는 캐글에서 어떤 것을 해야하는 지 잘 모르는 newbie 라면, 타이타닉을 하시는 게 가장 좋은 선택입니다.
 이 페이지를 5번 방문했습니다. 최근 방문 날짜: 19. 9. 5

Kaggle Titanic: Machine Learning model (top 7%) - Towards ...
<https://towardsdatascience.com/kaggle-titanic-machine-...> > 이 페이지 번역하기
 2018. 11. 4. - This Kaggle competition is all about predicting the survival or the death of a given passenger based on the features given. This machine ...

kaggle titanic

카글 필터

캐글 (Kaggle) 타이타닉 데이터 분석

11:38

Predicting Titanic Survivors With Machine Learning

Coding Tech · 조회수 3,2만회 · 1년 전

What's a better way to understand machine learning than a practical example? And who hasn't watched the classic 1997 movie?

Kaggle - Titanic Solution [1/3] - data analysis

Minsuk Heo 허민석 · 조회수 5만회 · 1년 전

[github source link] <https://github.com/minsuk-heo/kaggle-titanic/tree/master> This short video will cover how to define ...

캐글 타이타닉 Titanic - 1. Dataset check

You Han Lee · 조회수 4,8천회 · 1년 전

<http://kaggle-kr.tistory.com/> 가시면 감히자로 있어요! 캐글에 꼭 확인 대학원생입니다. 캐글을 시작하면 가장 먼저 하게되는 titanic 을 ...

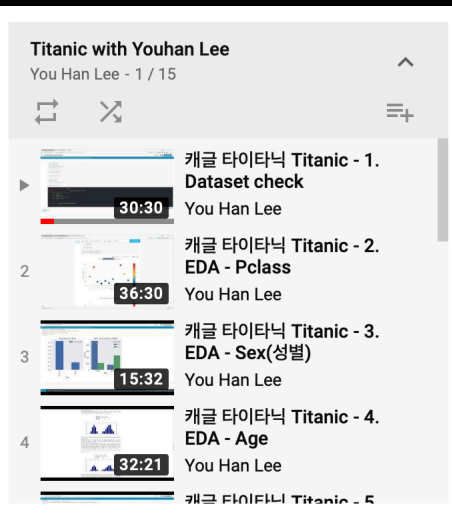
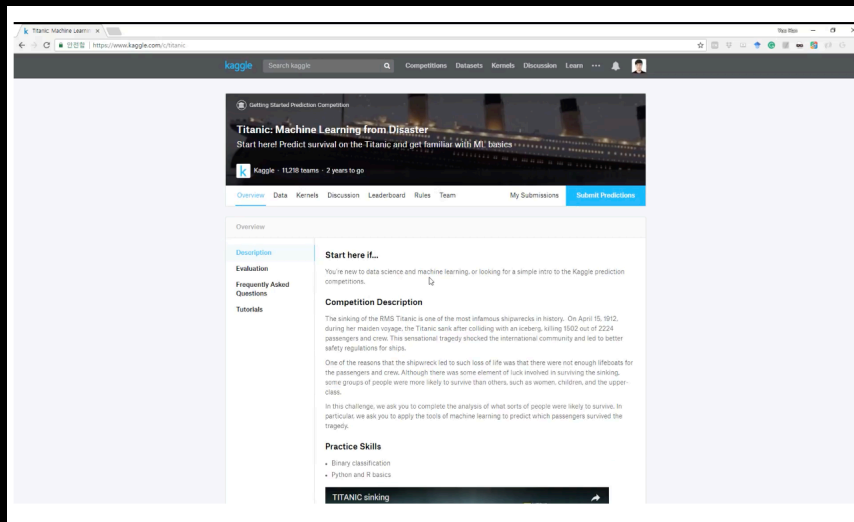
How to Get Started with Kaggle's Titanic Competition | Kaggle

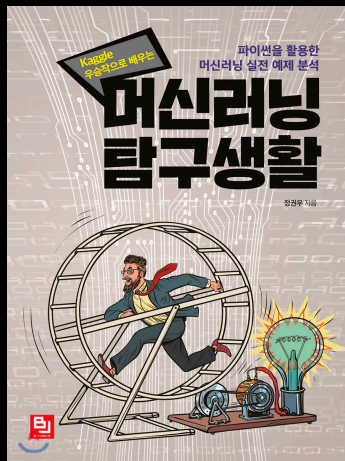
Kaggle · 조회수 6.5만회 · 2주 전

New to Kaggle? Our Titanic competition is a great place to start. In this video, Kaggle data scientist Dr. Rachael Tatman walks you ...

How to do the Titanic Kaggle competition in R - Part 1

Data Science Online · 조회수 6,1만회 · 2년 전





바로 시작!



SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019



THANK YOU

hwk06023@gmail.com / 010-3132-5914

SOSCON 2019

SAMSUNG OPEN SOURCE CONFERENCE 2019

