

한국의 캐글 그랜드 마스터

한국의 캐글 그랜드 마스터에 대해

2024.04.11

세종대학교 센B204

기획자 : 정재석

Who is kaglle?

캐글 첫걸음

스마클 정재석
E. wjdwtjr0318@gmail.com



Kaggle Progression System

Kaggle Progression System은 kaggle의 Tier를 나타냅니다.
메달을 따고 다양한 대회를 참여를 통해 티어를 올릴 수 있습니다.
상위 사이언티스트 들은 대회 리더보드에 등록됩니다.



Novice

You've joined the community.

☒ Register!



Contributor

You've completed your profile, engaged with the community, and fully explored Kaggle's platform.

- ☒ Run 1 notebook or script
- ☐ Make 1 competition or task submission
- ☐ Make 1 comment
- ☒ Give 1 upvote



Expert

You've completed a significant body of work on Kaggle in one or more categories of expertise. Once you've reached the expert tier for a category, you will be entered into the site wide Kaggle Ranking for that category.

Competitions

☐ 2 bronze medals

Datasets

☐ 3 bronze medals

Notebooks

☐ 5 bronze medals

Discussions

☐ 50 bronze medals



Master

You've demonstrated excellence in one or more categories of expertise on Kaggle to reach this prestigious tier. Masters in the Competitions category are eligible for exclusive Master-Only competitions.

Competitions

☐ 1 gold medal

☐ 2 silver medals

Datasets

☐ 1 gold medal

☐ 4 silver medals

Notebooks

☐ 10 silver medals

Discussions

☐ 50 silver medals

☐ 200 medals in total



Grandmaster

You've consistently demonstrated outstanding performance in one or more categories of expertise on Kaggle to reach this pinnacle tier. You're the best of the best.

Competitions

☐ 5 gold medals

☐ Solo gold medal

Datasets

☐ 5 gold medals

☐ 5 silver medals

Notebooks

☐ 15 gold medals

Discussions

☐ 50 gold medals

☐ 500 medals in total



About Rank!!

캐글의 등급을 알아보자

Contributor



1. 1번 이상의 notebook or script 실행
2. 1번 이상의 comptetion or task 제출
3. 1번 이상의 댓글 작성
4. 1번 이상의 upvote 하기

Expert



1. competitions 에서 2개 이상의 브론즈 매달
2. Datasets 에서 2개 이상의 브론즈 매달
3. Notebooks 에서 2개 이상의 브론즈 매달
4. Discussions에서 2개 이상의 브론즈 매달

Master



1. competitions 에서 1개 gold ,2 silver 매달
2. Datasets 에서 1개 gold ,2 silver 매달
3. Notebooks 에서 10개 이상의 silver 매달
4. Discussions에서 50개 이상의 silver medals , 200 medals i n total

Grandmaster



1. competitions 에서 5개 gold , solo gold medal
2. Datasets 에서 5개 gold ,5 silver 매달
3. Notebooks 에서 15개 이상의 gold 매달
4. Discussions에서 50개 이상의 gold medals , 500 medals i n total



김윤수

2021년까지 5명이던 한국인 그랜드마스터 6번째 그랜드마스터

- AI 비전공자 출신
- 그의 3가지 원동력! 열정, 시간, 그리고 훌륭한팀원

이번 대회 과제에 대해 기존 기술에서는 문장을 구성하는 단어 수, 단어의 문자 수, 언어의 품사 정보 등을 활용한 피처를 사용했으나 이는 텍스트 전체의 맥락을 파악하지 못하는 한계가 있었다. 업스테이지팀은 기존 통계적 피처에 텍스트 맥락을 파악하는 언어 모델을 추가한 ML 알고리즘을 개발해 성능을 한 차원 끌어올렸다.

출처 : AI타임스(<https://www.aitimes.com>)

캐글을 시작하며 도움되었던 3가지는?!

- 영어..! 구글링 그리고 코세라?
- 캐글은 빠르게 시작할 수록 좋다!!

'코세라'입니다. 온라인 강의 플랫폼인 코세라는 다양한 분야의 퀄리티 높은 강의를 제공하고 있습니다. 그중 앤드류 응 교수의 머신러닝과 딥러닝, 데이터 사이언스를 위한 파이썬을 수강했습니다. 영어로 진행되는 강의가 많은 만큼, 영어 청취가 가능하면 강의 선택지가 많아질 수 있을 것 같습니다.

김상훈



1년만의 캐글 그랜드마스터.. 엄청난 속도!!

- 스타트업 업스테이지 대표
- 김윤수 캐글 그랜드마스터와의 합작품

사람이 1위를 기록한 대회는 캐글이 주관하고 글로벌 온라인 쇼핑 플랫폼 쇼피(Shopee)가 주최한 '가격 일치 보장(Price Match Guarantee)'대회다.

온라인 쇼핑몰에서는 가격 비교를 위해 동일 상품을 그룹화하는 작업이 필요하다. 여기에 사용하는 AI 정확도를 높이는 것이 이번 대회 과제였다. 제대로 그룹화가 이뤄지지 못할 경우 사용자가 각 상품을 소개하는 상세 페이지에 일일이 들어가 같은 상품인지 확인해야 한다. 대회 참가자들은 상품 그룹화를 가능한 한 정확히 수행하는 AI를 개발해 이러한 고객 불편을 줄이는 미션을 수행했다.

가격 일치 보장 대회는 지난 3월부터 약 2개월 동안 진행됐다. 총 3만 달러(한화 3천 3백만 원)의 상금이 걸린 이 대회에는 전 세계에서 2464팀, 3079명 연구자가 참가했으며 업스테이지팀이 최고 성적을 거뒀다. 일본, 러시아 참가팀이 뒤를 이어 수상 명단에 올랐다.

출처 : AI타임스(<https://www.aitimes.com>)

이유한



카이스트의 원자력 박사 캐글 그랜드마스터가 되다

- 세계에서 196번째, 국내에서는 4번째다.
- 그의 3가지 원동력! 열정, 시간, 그리고 훌륭한 팀원

캐글이 주관하며 영구의 챔스가 주최한 AI 분자물성 예측 경진대회에서 그의 팀이 최종 3위를 기록 하여 당당히 금메달을 차지했다. 팀명은 "Solve chem. together"에서 알수있다 싶이 팀과 함께한다는 그의 뜻을 알 수 있다.

각 분야 전문가들과 아이디어를 모으다!

우수한 5명과의 팀합

최성한 박사, 송원호 학부생, 이영수 연구원, 김상훈 과 함께 대회에 나가 각자의 역할을 맡아 대회에 임하였다. 최성한 박사는 문제해결에 필수적인 도메인 지식을 팀원과 나눴으며 김상훈 데이터 사이언티스트는 gnn 방식 아닌 bert 모델을 사용해보았다.

앞선 3명의 그랜드마스터를 통해 우리가 배울점

영어, 구글링, 그리고 코세라 유기적 연결

"영어"

AI 기술과 논문 코딩 관련 자료의 대부분이 영문이기에 영어 공부는 선택이 아닌 필수이다.

"구글링"

모르는 문제를 마주하면 구글 검색을 통해 더욱 빠르게 답을 찾아 나갈 수 있다.

구글은 최고의 선생님이었습니다.

이때 영문으로 검색하는 편이 더 많은 양질의 정보를 습득하는 데 도움이 되고,
코드 오류에 관한 답을 찾기 위한 더 많은 선택지를 찾아볼 수 있어 효율적입니다.

"코세라"

온라인 강의 플랫폼인 코세라는 다양한 분야의 퀄리티 높은 강의를 제공하고 있습니다.

그중 앤드류 응 교수의 머신러닝과 딥러닝, 데이터 사이언스를 위한 파이썬을 수강했습니다.

영어로 진행되는 강의를 많은 만큼, 영어 청취가 가능하면 강의 선택지가 많아질 수 있을 것 같습니다.

발표를 마치며

문과 출신에서 캐글 그랜드 마스터가 되기까지의 여정 영상

<https://www.content.upstage.ai/events/tech-talks/how-to-become-a-kaggle-grand-master>



한국의 모든 캐글 그랜드 마스터를 찾고자 했으나 유감스럽게도 3명만을 발견할 수 있었습니다. 이 3명은 모두 호기심을 갖고 개인적으로 시작하여, 결국 팀으로 발전하며 그랜드 마스터가 되었습니다.