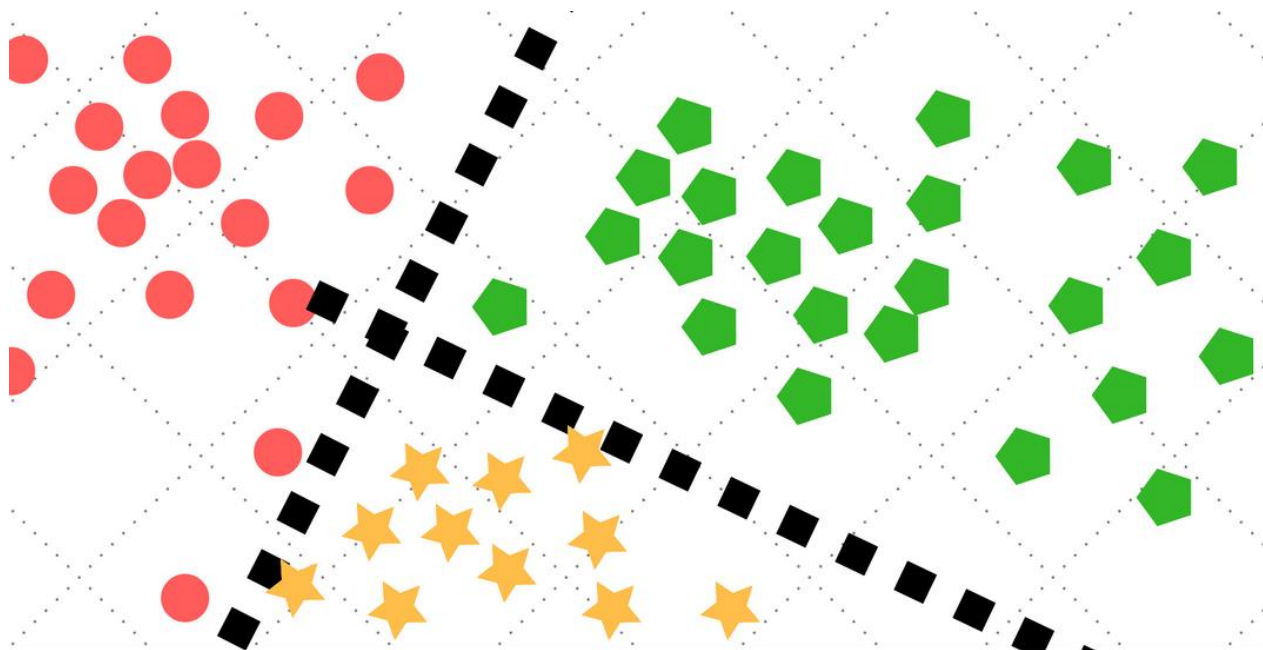


딥러닝/클라우드

2021 Classification 경진대회



참여 방법

- 제공되는 train_open.csv 파일을 이용하여 classification 모델 생성
(train_open.csv 데이터셋의 마지막 컬럼이 class label 임)
- 모델을 이용하여 testset.csv 파일의 자료에 대한 class 예측
(testset.csv 데이터셋에는 class label 이 없음. 따라서 참가자는 train_open.csv 만을 이용하여 모델 개발 및 평가를 해야함)
- 예측 결과 파일을 경진대회 사이트에 제출
- 경진대회 사이트는 예측결과를 정답과 비교하여 accuracy 계산
- 대회 종료 후 accuracy 에 의해 참가자의 순위를 결정
- 개인별로 30회까지 예측결과를 제출할 수 있음
- 제출한 예측 결과중 가장 정확도가 높은 것으로 최종 순위를 평가 받음
- 최종 평가시 제출기한이 지난것, 30회가 넘는 것은 평가에서 제외함

train_open.csv

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
1	3	529	5	4	4	3	1	0	5	1	3	2	2	3	5	1	32	0	2	3	3	2	1	
2	3	1846	3	3	3	4	2	10	3	1	5	3	2	3	2	2	22	12	2	1	3	3	1	
3	3	3240	5	4	5	4	1	0	4	1	5	5	1	4	1	2	61	0	2	5	4	1	2	
4	4	163	2	1	5	3	4	0	3	1	4	2	2	5	5	5	52	7	1	4	4	2	1	
5	3	2813	3	3	5	2	5	0	3	1	4	4	2	3	5	4	24	0	1	5	3	1	1	
6	4	587	3	5	4	5	4	0	5	1	3	3	2	4	1	0	55	0	1	3	3	1	2	
7	3	667	2	4	3	5	1	0	5	2	4	5	2	4	3	4	24	0	2	4	5	2	1	
8	4	145	5	3	4	0	3	0	2	1	2	1	2	1	1	0	30	0	1	5	5	2	2	
9	1	907	4	4	1	0	2	42	4	1	5	4	2	5	3	0	30	33	2	5	5	2	2	
0	3	184	5	3	4	1	3	30	3	1	1	3	2	3	3	4	57	14	1	5	5	2	1	
1	2	2388	2	2	2	3	1	0	4	1	2	2	2	2	2	1	15	7	1	2	4	1	1	
2	4	391	3	3	5	3	2	0	5	1	1	4	2	3	4	2	26	0	1	3	5	2	1	
3	2	198	4	3	5	4	3	42	4	1	3	4	2	4	4	5	23	65	2	4	4	2	1	

마지막 컬럼이 class label

test_open.csv

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
1	1	1262	5	3	5	4	3	0	3	1	4	1	2	5	3	1	53	0	1	3	4	3		
2	4	241	5	5	3	5	4	5	5	1	2	4	2	5	1	0	42	0	1	4	5	1		
3	2	1590	2	1	1	2	1	0	5	2	4	4	2	5	1	5	25	4	1	5	5	2		
4	3	517	4	4	5	2	3	0	5	1	3	2	2	3	1	2	29	0	1	2	1	2		
5	4	1649	5	5	1	5	4	0	4	1	3	3	1	5	4	5	38	0	1	2	4	1		
6	5	2330	4	4	2	2	3	66	1	1	3	1	2	4	4	3	13	55	1	3	3	1		
7	3	888	1	3	2	2	2	0	3	1	1	3	2	3	2	2	20	0	1	4	4	2		
8	5	2951	4	4	2	4	5	0	3	1	5	4	2	4	4	4	26	0	2	4	4	1		
9	5	1846	3	3	5	5	2	55	4	1	5	4	2	3	2	4	58	39	2	5	5	1		
10	1	997	2	4	4	1	2	0	3	1	3	1	2	4	4	1	22	1	2	1	1	1		

class label 없음

경진대회 사이트

<http://220.149.235.96:3838/competition/>

클릭하여
예측결과파일을
업로드

Machine learning competior 2021

Choose CSV File

Browse... 32 .csv

Upload complete

Ranking List

ranking	sid	acc	upcnt	date
1	32160699	0.95	2	2020-10-12
2	32160692	0.89	1	2020-10-12
3	test	0.75	1	2020-10-12
4	32160691	0.68	1	2020-10-12

Your ranking: 1 , acc: 0.948689956331878 , total upload: 2

My test log

sid	date	acc
2020-10-12	32160699	0.95
2020-10-12	32160699	0.95

현재까지의 정확도 순위
(대회기간중에는 테스트
데이터의 50%에 대한
accuracy만 평가하여 표시)

현재 나의 순위, acc,
총 업로드회수

나의 예측결과
제출 이력

예측결과 포맷

	A
1	answer
2	5
3	5
4	4
5	4
6	5
7	4
8	4
9	4
10	5
11	4
12	5
13	4
14	5
15	5

왼쪽과 같이 예측결과가
세로 방향으로 저장되어야 함
(예측 결과가 헤더 다음 행부터 시작)

파일 이름 포맷: 학번_이름.csv
예) 321601001_홍길동.csv

** 포맷이 정확하지 않으면 결과 업로드가 되지 않음

.csv 파일의 업로드시 오류가 발생하면 .xlsx 파일로 저장하여 재시도

기타사항

- 수업시간에 배운 것 포함하여 최대한 많은 classification 알고리즘을 테스트
(수업시간에 안 배운 알고리즘, 방법의 사용도 가능)
- Feature selection, hyper parameter tuning 등 배운 지식을 총동원하여 최상의 예측 모델을 만들기 바람

평가

- 제출결과의 정확도를 평가하여 점수 부여
- 미제출시 평가점수 없음
- 모델 도출 및 평가과정을 보고서로 작성하여 제출해야 함 (소스코드, 실행 결과도 포함)
- 보고서 마지막 부분에 경진대회 참여 소감을 서술할 것
- 보고서 평가는 배운 내용을 얼마나 충실히 반영 했는지, 좋은 모델을 만들기 위해 얼마나 노력했는지가 주된 기준임 (가급적 상세한 설명 바람)
- 경진대회 점수(60%) + 보고서 점수(40%)가 중간고사 점수임

GOOD LUCK!