

Thyroid Gene Classification User Manual

▶ Drag & Drop Data File(CSV or Excel)

- 분석하고 싶은 환자의 데이터를 아래와 같은 형태의 .csv or .excel로 만들어 넣어주세요.

Name	gene 1	gene 2	gene 3	...	gene n
sample 1	2.3121	4.1431	10.4535	...	3.3290
sample 2	5.3424	5.6345	4.34234	...	3.6238
sample 3	7.3423	8.7674	12.3456	...	4.2450
sample 4	2.3434	3.3734	7.43278	...	5.7888
...	5.8653
sample n	4.4353	4.6865	32.77854		3.2356

- 환자 한명의 gene 발현량정보는 한 행(row)에 저장되어 있어야 합니다.

- 데이터에 결측치가 없어야 합니다.(결측치란 숫자 0이 아닌 None)

- gene name은 official gene symbol이어야 합니다.

* **note** : test dataset은 24.02.22에 제공받은 "Thyroid_GSE60542_EXP_20638genes_92cases" 데이터를 기준으로 전처리 및 학습이 되었기에, test 진행시 **기 제공받은 데이터의 feature와 동일**해야 분석가능.

(만약 **error발생시**, test하고자 하는 **dataset**을 전달해주시면 전처리 및 test web update를 진행하도록 하겠습니다.)

▶ Model Weight Path

- 학습된 모델의 weight file이 있는 경로입니다.

▶ Result CSV Path

- 예측 결과가 저장될 파일 경로 입니다.
- 예측 결과는 다음과 같은 형태로 저장 됩니다.

	predict proba	predict value
sample 1	0.9934523	N1
sample 2	0.3245517	N0
...
sample n	0.1223333	N0