

자연어처리_과제#8

1. 강의자료 16-B장의 4쪽에서 설명한 바와 같이 Hugging face transformers를 설치한다. 다음 text 데이터를 이용하여 pipeline 방식의 텍스트 분류, 개체명 인식, 질문 답변, 요약을 차례로 수행하고 코드와 결과를 제시하라.

```
!pip install transformers
```

```
# 주어진 텍스트 데이터
text = """The South Korean general appointed as the commander of President Yoon Suk Yeol's short-lived declaration of martial law learned of the move only when Mr. Yoon announced it on television, the general told lawmakers during hearings on Thursday. Despite his position, the general, Park An-su, told a parliamentary committee that he also did not know who ordered troops to move in on the National Assembly in an attempt to cordon it off. During his testimony, the deputy defense minister, Kim Seon-ho, said that the defense minister had ordered the troops in. Mr. Kim said that he had opposed the mobilization. The defense minister, Kim Yong-hyun, resigned Thursday before the hearings and did not take part. The hearings were an inquiry into why martial law was declared for several hours on Tuesday night. Mr. Yoon's extraordinary move plunged the country into a political crisis, caused widesp"""
```

텍스트 분류

```
from transformers import pipeline

classifier = pipeline("text-classification")

outputs = classifier(text)
print(outputs)
```

```
[{'label': 'NEGATIVE', 'score': 0.9855273962020874}]
```

개체명 인식

```
import pandas as pd
ner_tagger = pipeline("ner", aggregation_strategy="simple")

outputs = ner_tagger(text)
pd.DataFrame(outputs)
```

	entity_group	score	word	start	end
0	MISC	0.999056	South Korean	4	16
1	PER	0.896835	Yoon Suk Yeol	65	78
2	PER	0.984267	Yoon	154	158
3	PER	0.993398	Park An	278	285
4	PER	0.879135	su	286	288
5	ORG	0.993673	National Assembly	384	401
6	PER	0.947953	Kim Seon	485	493
7	PER	0.741048	ho	494	496
8	PER	0.998273	Kim	560	563
9	PER	0.996000	Kim Yong	629	637
10	PER	0.793684	hyun	638	642
11	PER	0.980931	Yoon	808	812

질문 답변

```
reader = pipeline("question-answering")
question = "Who is Kim Yong-hyun?"
outputs = reader(question=question, context=text)
pd.DataFrame([outputs])
```

	score	start	end	answer
0	0.610437	607	627	The defense minister

요약

```
summarizer = pipeline("summarization")
outputs = summarizer(text, max_length=45,
clean_up_tokenization_spaces=True)
print(outputs[0]['summary_text'])
```

President Yoon Suk Yeol’s short-lived declaration of martial law only learned of the move only when he announced it on television. Despite his position, the general told a parliamentary committee that he also