

# NewSeconds

뉴세컨즈

팀 COPI

201601620 이창민

201702994 이정민

201801619 송채원

201801623 윤세휘



## CONTENTS

- 01 프로젝트 소개
- 02 프로젝트 필요성
- 03 프로젝트 구조
- 04 프로젝트 주요 기능
- 05 사용 기술
- 06 활용 방안 및 기대효과
- 07 시연 영상



# Part 1.

---

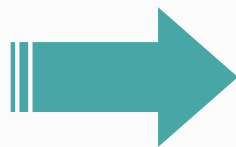
## 프로젝트 소개



# 뉴세컨즈(NewSeconds)

"인터넷뉴스를 자동으로 요약하여 카드뉴스로 보여주고 기사를 추천해준다"

[이데일리 조웅석 기자] 22일 코스피 지수가 하락 마감하며 3150선 아래로 떨어졌다. 장내내 개인 순매수와 기관의 순매도가 팽팽하게 맞선 가운데 외국인의 수급에 따라 보험권에서 등락을 반복했다.마켓포인트에 따르면 이날 코스피 지수는 전 거래일 대비 0.64%(20.21포인트) 하락한 3140.63에 마감했다. 전일 종가기준 사상 최고치를 경신하는 등 최근 3거래일 연속 상승했던 코스피 지수는 다소 숨을 고르는 모양새다.개인이 1조6121억원 순매수, 기관이 1조3711억원을 순매도하면서 힘겨루기를 하는 동안 외국인 수급에 따라 지수가 움직였다. 1000억원대 순매도 시작한 외국인이 오전 11시께 순매도 규모를 200억원대로 줄이면서 코스피 지수는 상승 전환했으나 이후 다시 매도물량을 늘리자 함께 하락했다. 외국인은 결국 2659억원 순매도로 마감했다. 기관에서는 금융투자가 6697억원으로 가장 순매도 규모가 컸다.프로그램 매매는 차익과 비차익을 더해 7218억원 순매도였다. 원-달러 환율은 5.0원 상승한 1103.20에 거래를 마쳤다.시가총액 상위는 하락 우위였다. 대장주인 삼성전자(005930)가 1%대 하락한 것을 포함해 SK하이닉스(000660), LG화학(051910), 현대차(005380), 삼성바이오로직스(207940), 셀트리온(068270), 기아차(000270), 현대모비스(012330), LG전자(066570), 삼성물산(028260) 등도 모두 하락 마감했다. 반면 외국인 순매수금액 1,2위인 NAYER(035420)와 삼성SDI(006400)는 6%대의 높은 상승세를 보였다.업종별로도 하락이 다소 많았다. △운수장비 △철강 및 금속 △금융업 △보험 등이 2%대 하락했고 △증권 △전기가스업 △은행 △유통업 등도 1%대 하락했다. △반면 비금속 광물은 4%대, 섬유 의복은 3%대 올랐고 이외에도 △서비스업 △음식료품 △종이,목재 등도 상승 마감했다.종목별로는 수소 관련주인 두산퓨얼셀1우(33626K)가 가격 제한폭까지 올랐고 두산퓨얼셀2우B(33626L)도 23% 넘게 상승했다. 조 바이든 미국 대통령 취임 후 친환경 관련주에 대한 기대감이 반영된 것으로 보인다. 태경케미컬(006890)도 26% 넘게 상승 마감했다.이날 거래량은 10억9424만2000주, 거래대금은 23조9627억3200만원으로 집계됐다. 상한가 1개 종목을 포함해 383개 종목이 올랐고 하한가 없이 476개 종목이 하락했다. 47개 종목은 보험권에서 거래를 마쳤다.한편 21일(현지시간) 뉴욕증권거래소(NYSE)에서 다우존스 30 산업평균지수는 전장보다 0.04%(12.37포인트)하락한 3만1176.01에 마감했다. 반면 스탠더드앤드푸어스(S&P) 500 지수는 전장보다 0.03%, 기술주 중심의 나스닥 지수는 0.55% 오른 채 장을 마쳤다. S&P500·나스닥 지수는 전일에 이어 다시 증가 기준 최고가를 경신했다.조웅석 (chojuri@edaily.co.kr)

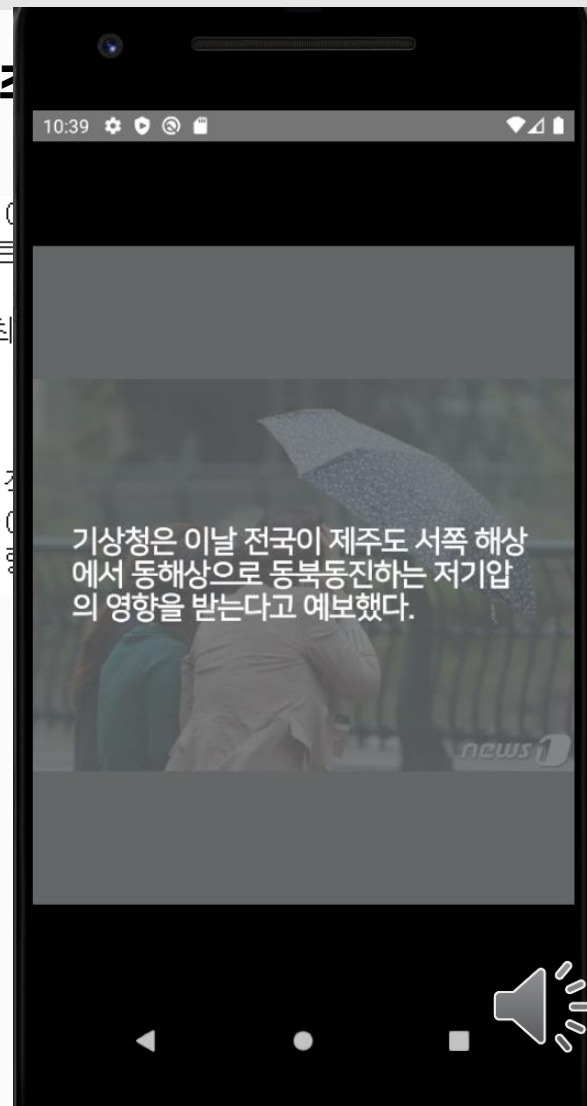


마켓 포인트에 따르면 코스피 지수는 전일 종가 대비 0.64% (20.21 포인트) 하락한 3140.63에 마감했다.

전일 종가기준 사상 최고치를 경신하는 등 최근 3거래일 연속 상승했던 코스피 지수는 다소 숨을 고르는 모양새다.

1000억원대 순매도 시작한 외국인이 오전 11시께 순매도 규모를 200억원대로 줄이면서 코스피 지수는 상승 전환했으나 이후 다시 매도물량을 늘리자 함께 하락했다.

기상청은 이날 전국이 제주도 서쪽 해상에서 동해상으로 동북동진하는 저기압의 영향을 받는다고 예보했다.



# Part 2.

---

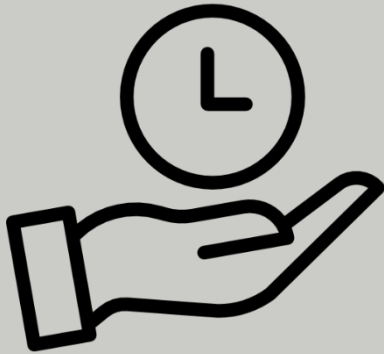
## 프로젝트 필요성



## "NewSeconds" 프로젝트의 필요성



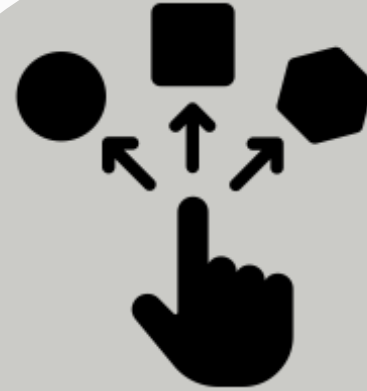
## "NewSeconds" 프로젝트의 필요성



효율적인 기사 읽기



개인화



편향되지 않은  
다양한 콘텐츠 제공



# Part 3.

---

## 프로젝트 구조







Amazon  
EC2

# 프로젝트 구조



Amazon  
RDS

AWS EC2



python™

Crawling  
텍스트 요약

AWS RDS



MySQL

추천시스템

1 • `select * from NEWS1000;`



Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Co

TITLE	CONTENT1	CONTENT2	CONTENT3	CONTENT4	CONTENT5	CONTENT6	CONTENT7
주호영...	국민의힘 ...	뉴스1국민...	주 권한대...	주 권한대행...	새 원내대...	주 권한대...	안 대표도 ...
대권후...	/자료제공 ...	엠브레인...	차기 대통...	이 지사는 ...	서울과 대...	연령별로 ...	특히 40대...
'초과이...	심상정 정...	2021.4.29...	그는 토초...	그러면서 "...	우선 "개인...	"며 "단, 토...	"며 "부동...
문자 폭...	더불어민주...	강성 지지...	사진=연합...	윤 의원은 "...	김용민 의...	조용천 의...	문자를 보...
"반도체...	"반도체 대...	28일 서울 ...	업계가 가...	이날 주제 ...	박 회장은 ...	또 삼성전...	당장 확보...
진중권 ...	"민주당은 ...	진 전 교수...	진 전 교수...	진 전 교수...	이 지사에 ...	진중권 전 ...	당대표 후...

RDS > 데이터베이스 > newdb

newdb

요약

DB 식별자  
newdb

EC2 > 인스턴스 > i-0487ee3c09751e29d

i-0487ee3c09751e29d (2번)에 대한 인스턴스 요약

less than a minute 전에 업데이트됨

인스턴스 ID

i-0487ee3c09751e29d (2번)

인스턴스 상태

실행 중

"민주당... 2021.04.29. 오전 9:07 100 6 https://imnews.pstatic.n




# Part 4.

---

## 프로젝트 주요 기능



## 주요 기능

-  **01. 텍스트 요약**  
포털 사이트들의 기사 텍스트에 대하여 **추출 요약**을 수행함.
-  **02. 관련 기사 추천**  
사용자가 **최근 (20초 이상) 읽은 기사 3개와 가장 유사한 기사**를 각 기사와 같은 카테고리 내에서 5개씩 총 15개 추천한다.
-  **03. 카드뉴스 구성**  
요약한 문장들과 기사의 이미지를 활용하여 사용자가 보기 편리하게 카드뉴스를 구성한다.



# Part 5.

---

사용기술



# 사용 기술



## 01. 크롤링

- BeautifulSoup
- Selenium



## 02. 텍스트 요약

- TextRank 알고리즘
- Tf-Idf Vectorizer
- 자연어처리 라이브러리 활용 (konlpy등)



## 03. 관련 기사 추천

- contents 기반 추천시스템
- Cosine 유사도 계산
- Tf-Idf Vectorizer
- pandas, 사이킷런 라이브러리 등 사용



## 04. 사용자 인터페이스 구성

- Fragment Activity
- Recycler View



## 05. MySQL 기반 데이터베이스 시스템 구축

- AWS RDS에서 생성
- 파이썬, Flask와 상호작용



## 06. AWS EC2 서버 구축

- Flask 웹서버 구동
- 파이썬 코드 실행



## 텍스트 요약



### 추출형 문서 요약

- 본문에서 **의미있는 부분을 선택하고 추출**하여 요약하는 방법이다.
- 일종의 하이라이터 역할.



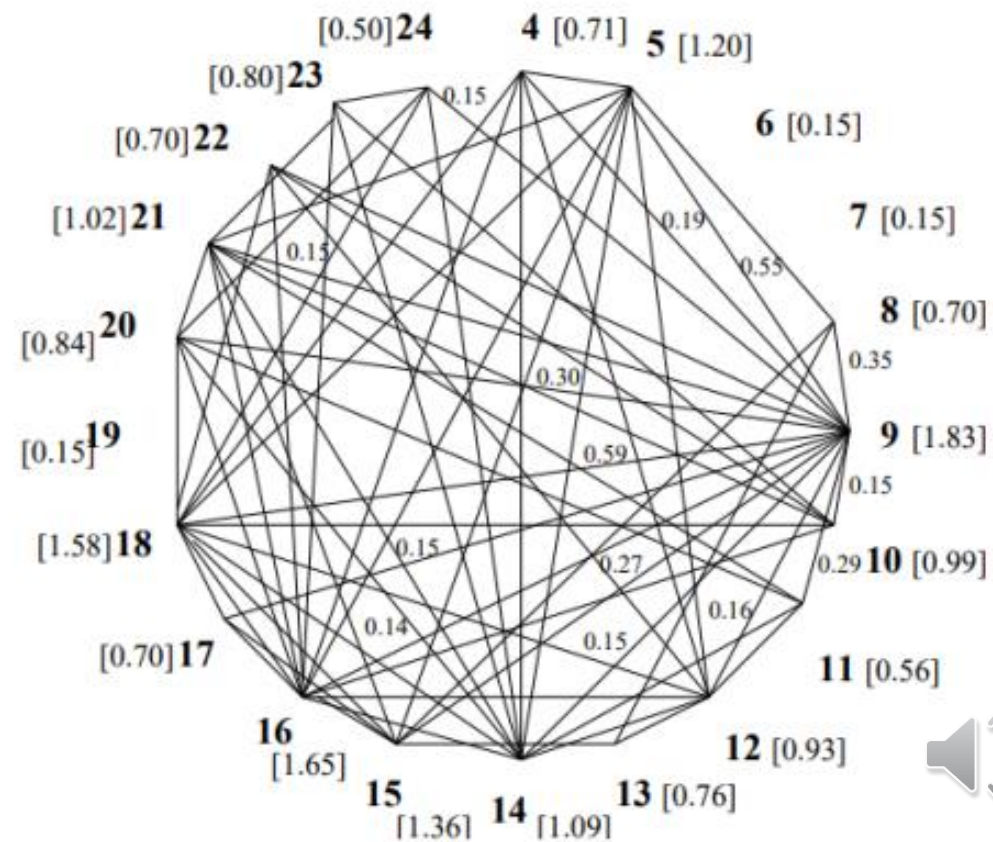
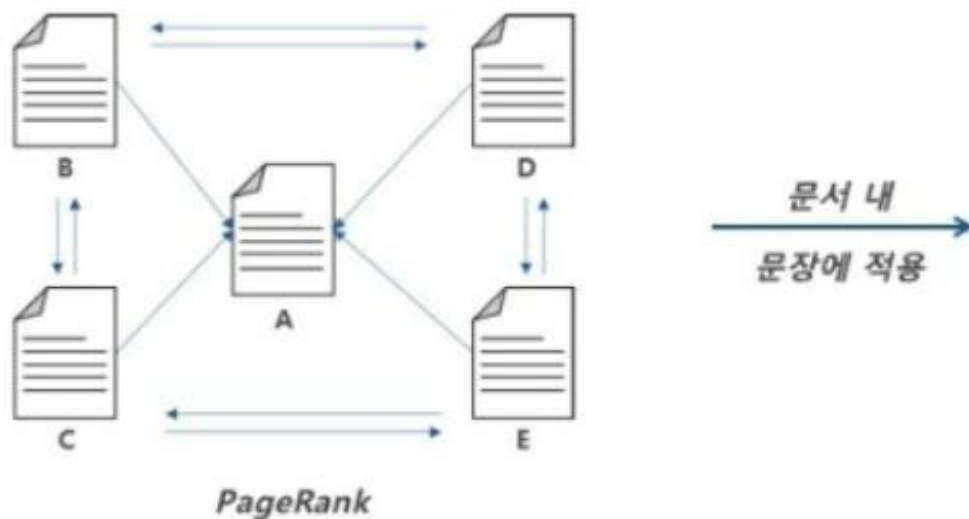
### 생성형 문서 요약

- RNN과 같은 자연어 생성 기술을 활용하여 **기존 텍스트 문장을 토대로 창의적인 새로운 문장을 작성**하는 방법이다.
- 학습시키기 위해 필요한 원문 텍스트와 요약된 텍스트가 필요하다.



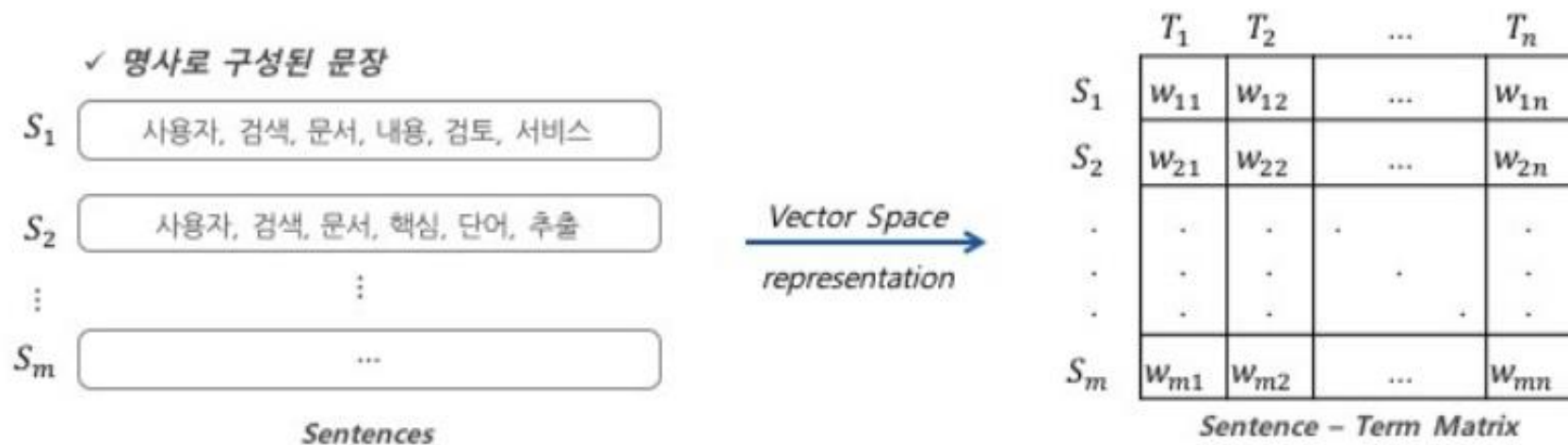
# 텍스트 요약: TextRank 알고리즘

- 'TextRank'는 텍스트에 관한 **graph-based ranking model**로서 노드들의 관계를 중심으로 요약문을 작성한다. Google의 PageRank의 영향을 받은 방법이다.
- 'PageRank'는 구글 검색의 기초가 되는 방법으로 **각 인터넷 페이지**에 **PageRank 점수**를 부여한다. TextRank는 이를 **문장에 적용**한 방법이다.

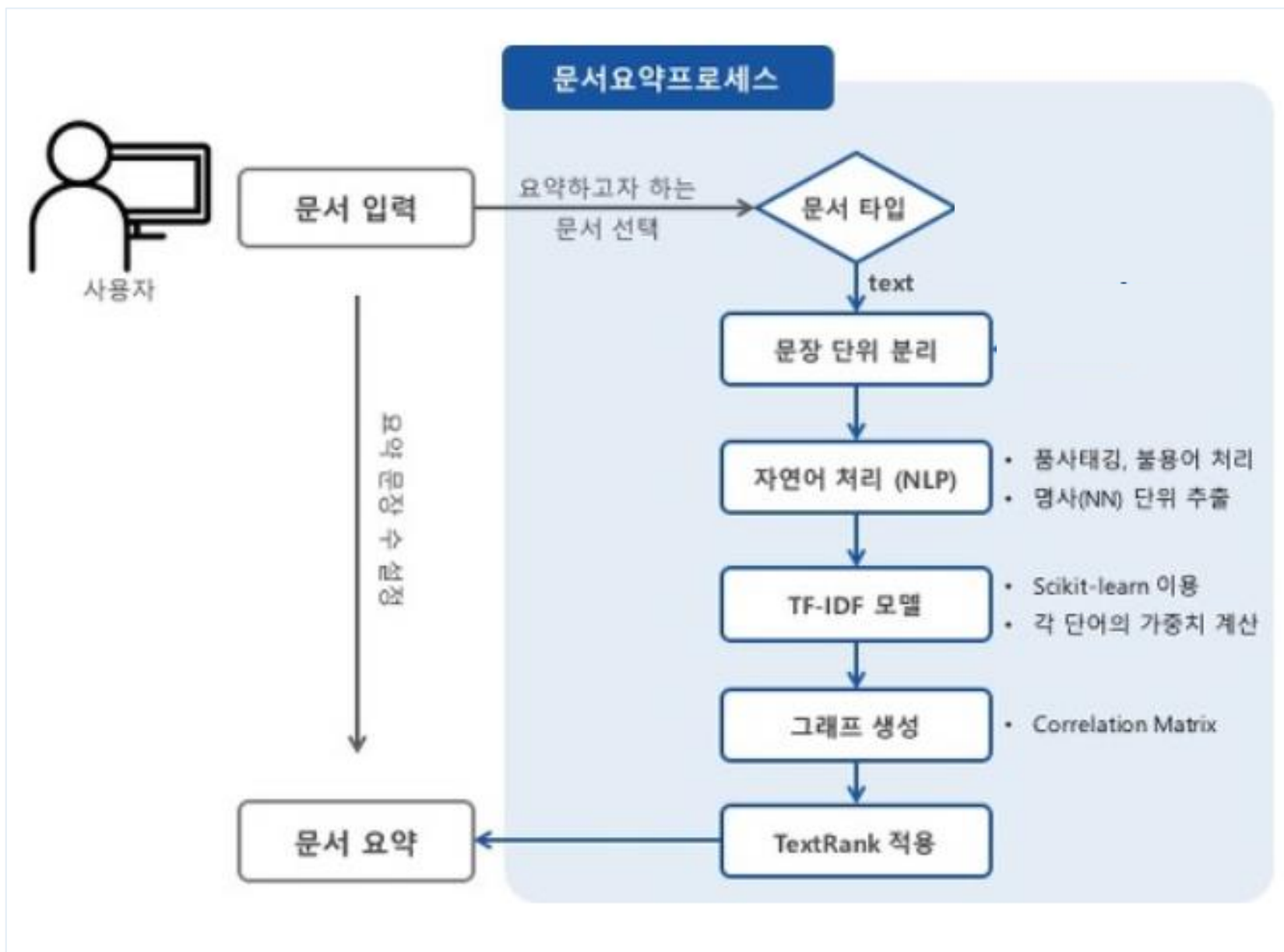


## 텍스트요약: Tf-Idf

- **TF-IDF**(Term Frequency - Inverse Document Frequency)는 **각 문장 사이의 유사성을 계산**하는 방법.
  - 한 문서에서 특정 단어의 출현빈도가 높다면 중요한 단어이다.
  - 다른 문서에도 많이 등장하는 일반적인 단어라면 중요도는 낮아진다.
- **TF(Term Frequency)** : 단어 빈도. 특정 단어가 문서 내에 얼마만큼의 빈도로 등장하는지를 나타내는 척도
- IDF(Inverse Document Frequency)** : 한 단어가 전체 문서에서 얼마나 공통적으로 등장하는지를 나타내는 척도







# 기사 요약 실행 결과 1

## [제목]

일론 머스크, 비트코인 버리고 도지코인 갈아타나

## [본문]

[파이낸셜뉴스] 테슬라 결제수단에서 비트코인을 제외한다고 발표해 가상자산 시장에 충격을 준 일론 머스크 테슬라 CEO가 이번에는 도지코인(DOGE)과 협력 작업을 하고 있다는 트윗을 남겼다. 머스크 트윗에 도지코인은 한때 20% 이상 급등하는 등 여지없이 반응했다. 투자자들은 일론 머스크의 행동이 시장 변동성을 높이고 있는 것에 분노하면서도 도지코인이 테슬라 결제수단이 될 수 있을지 관심을 기울이고 있다. 일론 머스크는 13일(현지시간) "시스템의 거래 효율성을 개선하기 위해 도지코인 개발자들과 함께 일하고 있다"며 "잠재적으로 유망하다"고 썼다. 도지코인은 전날 0.39달러까지 떨어졌지만 머스크의 트윗 이후 0.49달러(25.6%)까지 수직 상승했다. 도지코인이 조만간 미국 대형 거래소 바이낸스에 상장된다는 소식도 상승에 일조했다. 도지코인은 오히려 머스크의 일거수 일투족에 따라 급등락을 반복해왔다. 머스크가 미국 예능프로그램 SNL(Saturday Night Live)에 출연한 이후에는 30% 가까이 급락했다가 급반등했다. 전날 테슬라 결제수단에서 비트코인을 제외한다는 발표에 가상자산 시장 전체가 충격에 빠지며 또 다시 급락하기도 했다. 투자자들은 일론 머스크의 행동에 "작전 아니냐"며 불만을 털어놓고 있다. 앞서 머스크는 "비트코인 채굴과 거래에 화석연료 사용이 급격하게 증가하는 것에 우려를 표한다"며 테슬라 결제 수단에서 비트코인을 제외한다고 밝혔다. 머스크는 '본인을 도지코인의 아버지'(Dogefater)라고 부르며 달까지 도지코인을 가져가겠다고 트윗하는 등 도지코인의 열성적 지지자다. 비트코인이 전력 소모가 크다는 지적은 끊이지 않아왔다. 케임브리지 연구진은 지난 2월 가상자산 채굴에 들어가는 전력이 연간 121.36테라와트시(TWh)라는 추정을 내놨다. 아르헨티나(121TWh) 네덜란드(108.8TWh) 아랍에미리트(113.20TWh)의 연간 전력 소비량을 넘어서는 것이다. 연구진은 "비트코인 가격이 상승하면 더 많은 채굴자들이 더 많은 컴퓨터를 사용해 채굴에 나서며 전기 사용량 역시 늘어날 것"이라고 전망했다. 머스크는 특히 석탄에 대해 "다른 어떤 연료보다 많은 배출물을 내놓는다"고 지적했다. 현재는 중국 신장과 내몽고 등 석탄 광산이 많은 지역에서 채굴이 진행되는 경우가 많다. 풍부한 석탄을 이용한 전기를 싸게 공급받아 채굴에 나서는 것이다. 환경 이슈와 함께 미국과 중국이 대립각을 세우고 있는 분위기가 반영된 발표인 것 아니냐는 관측도 나온다. 석유 매장이 많은 이란도 미국의 경제제재를 피하기 위해 비트코인 채굴에 전격적으로 나서고 있다. 채굴은 복잡한 수식을 컴퓨터를 통해 풀고 그에 대한 보상으로 소량의 가상자산을 받는 형식이다. 비트코인의 경우 총 발행 한도가 제한돼 있어 시간이 지날수록 수식의 난이도가 높아진다. 더 많은 컴퓨터를 이용해 더 오랜 시간 계산을 해야 채굴에 성공할 수 있다. 채굴은 가상자산 생태계에서 거래 위변조를 막기 위한 필수 작업이다. 이날 급등은 도지코인이 새로운 테슬라 결제수단이 될 수 있을지에 대한 기대감이 반영된 것으로 보인다. 머스크는 전날 발표에서 "거래처리에 들어가는 에너지 소모량이 비트코인의 1% 미만인 다른 가상자산을 검토 중이다"라고 썼다. 채굴에 필요한 전력량은 도지코인이 압도적으로 낮은 것으로 알려져 있다. 미국 텍사스 TRG 데이터센터의 최근 자료에 따르면 비트코인의 거래처리(transaction)당 소비전력량은 707Kwh이다. 다른 코인들은 이더리움(62.56) 비트코인캐시(18.957) 라이트코인(18.522) 등의 순이다. 반면 도지코인은 0.12Kwh 수준이다. 발행량 제한이 없어 상대적으로 덜 복잡한 암호를 풀게 돼 있어 전력소모가 적다고 외신들은 보도했다. #비트코인 #가상자산 #도지코인 #전력소모량



## [문장1]

[파이낸셜뉴스] 테슬라 결제수단에서 비트코인을 제외한다고 발표해 가상자산 시장에 충격을 준 일론 머스크 테슬라 CEO가 이번에는 도지코인(DOGE)과 협력 작업을 하고 있다는 트윗을 남겼다.

## [문장2]

투자자들은 일론 머스크의 행동이 시장 변동성을 높이고 있는 것에 분노하면서도 도지코인이 테슬라 결제수단이 될 수 있을지 관심을 기울이고 있다.

## [문장3]

도지코인은 전날 0.39달러까지 떨어졌지만 머스크의 트윗 이후 0.49달러(25.6%)까지 수직 상승했다.

## [문장4]

전날 테슬라 결제수단에서 비트코인을 제외한다는 발표에 가상자산 시장 전체가 충격에 빠지며 또 다시 급락하기도 했다.

## [문장5]

앞서 머스크는 "비트코인 채굴과 거래에 화석연료 사용이 급격하게 증가하는 것에 우려를 표한다"며 테슬라 결제 수단에서 비트코인을 제외한다고 밝혔다.

## [문장6]

머스크는 '본인을 도지코인의 아버지'(Dogefater)라고 부르며 달까지 도지코인을 가져가겠다고 트윗하는 등 도지코인의 열성적 지지자다.

## [문장7]

채굴에 필요한 전력량은 도지코인이 압도적으로 낮은 것으로 알려져 있다.



## 기사 요약 실행 결과 2

### [제목]

경찰 “한강 사망 의대생, 부검 결과 익사 추정”

### [본문]

경찰이 한강에서 숨진 채 발견된 의대생 고(故) 손정민(22)씨의 사인을 익사로 추정했다. 서울경찰청 관계자는 13일 “전날 국립과학수사연구소(국과수)가 ‘사인은 익사로 추정되고, 머리의 두 군데 좌열창(挫裂創)은 사인으로 고려할 정도로 보기 어렵다’는 내용의 부검 감정서를 보내왔다”며 “부검 결과와 관계없이 모든 가능성을 열어놓고 수사에 총력을 다 하겠다”고 밝혔다. 정민씨 아버지인 손현(50)씨는 지난 1일 “정민씨의 왼쪽 귀 뒷부분에 손가락 2마디 크기의 자상이 2개 있다”고 말했다. 경찰은 정민씨와 서울 반포 한강공원에서 함께 술을 마셨던 친구 A씨와 9명의 목격자를 조사하고 카드 사용 내역도 확인했다. 조사 결과 정민씨와 A씨는 지난달 24일 오후 10시54분부터 다음날 오전 1시31분까지 인근 편의점에서 총 3차례에 걸쳐 술 9병을 구입했다. 목격자 진술에 따르면 A씨는 간혹 강 쪽으로 가서 토하고 오기를 반복했고, 정민씨는 대체로 누워있거나 앉아있는 모습이었다. 목격자들은 공통적으로 “그날 오전 2시부터 3시38분까지는 정민씨와 A씨가 함께 있었다”고 진술했다. 서울청 관계자는 “A씨가 오전 3시37분경 통화를 하고 있었고 그 옆에 정민씨가 앉아 있었다는 목격자 진술이 있다”고 말했다. 또 다른 목격자는 “그날 오전 4시20분쯤 A씨가 강에 가까운 잔디 끝 경사면에 누워 잠들어 있는걸 발견하고 위험하다는 생각이 들어 깨웠다. 하지만 당시 정민씨를 보지는 못했다”고 진술했다. A씨는 정민씨와 돛자리를 깔고 앉았던 장소에서 강쪽으로 10여m 이동한 곳에서 잠든 채 발견됐다. 경찰은 이에 따라 오전 3시38분~4시20분까지 40여분 동안의 두 사람 행적을 재구성하는 데 수사력을 집중하고 있다. 국과수는 부검 보고서에 정민씨가 마지막 음주 후 2~3시간 안에 사망했을 거라는 추정을 담았다. 이에 대해 경찰 관계자는 “음주 후 비교적 짧은 시간 안에 사망했을 거라는 취지”라며 “반드시 2~3시간 이내에 사망했을 거라는 뜻은 아니다”라고 설명했다. 정민씨의 혈중알코올농도에 대해선 “부검 결과 확인이 됐지만, 가족에게만 통보한 상태”라고 밝혔다. “두 사람이 평소 자주 만나는 사이였느냐”는 질문에는 “이 부분은 평가가 조심스럽다”고 답했다. 다만 두 사람이 다투는 모습을 봤다는 내용의 목격자 진술은 없었다는 게 경찰의 설명이다.



### [문장1]

경찰이 한강에서 숨진 채 발견된 의대생 고(故) 손정민(22)씨의 사인을 익사로 추정했다.

### [문장2]

서울경찰청 관계자는 13일 “전날 국립과학수사연구소(국과수)가 ‘사인은 익사로 추정되고, 머리의 두 군데 좌열창(挫裂創)은 사인으로 고려할 정도로 보기 어렵다’는 내용의 부검 감정서를 보내왔다”며 “부검 결과와 관계없이 모든 가능성을 열어놓고 수사에 총력을 다 하겠다”고 밝혔다.

### [문장3]

경찰은 정민씨와 서울 반포한강공원에서 함께 술을 마셨던 친구 A씨와 9명의 목격자를 조사하고 카드 사용 내역도 확인했다.

### [문장4]

목격자들은 공통적으로 “그날 오전 2시부터 3시38분까지는 정민씨와 A씨가 함께 있었다”고 진술했다.

### [문장5]

서울청 관계자는 “A씨가 오전 3시37분경 통화를 하고 있었고 그 옆에 정민씨가 앉아 있었다는 목격자 진술이 있다”고 말했다.

### [문장6]

또 다른 목격자는 “그날 오전 4시20분쯤 A씨가 강에 가까운 잔디 끝 경사면에 누워 잠들어 있는걸 발견하고 위험하다는 생각이 들어 깨웠다. 하지만 당시 정민씨를 보지는 못했다”고 진술했다.

### [문장7]

국과수는 부검 보고서에 정민씨가 마지막 음주 후 2~3시간 안에 사망했을 거라는 추정을 담았다.



# 기사 추천 실행 과정 및 결과 1

cosine\_sim함수의 파라미터는 데이터베이스에 들어있는 기사의 count

추천 목적 기사:

count	title	category	content
53	아파트 공시가격 기초자료 나왔지만...여전히 부족한 설명	300	(서울=연합뉴스) 윤종석 홍국기 기자 = 아파트 등 공동주택의 공시...

추천 가능 기사:  
(dataframe)

count	title	category	#
0	1	주호영·안철수 "합당 공감...큰 어려움 없을 것"	100
1	2	대권후보 적합도 이재명 24%vs윤석열 23%...이낙연은 또 한자릿수	100
2	3	'초과이득 절반 세금으로'...심상정, 폐지 23년 만에 토초세법 발의	100
3	4	문자 폭탄이 뒤길래...강경파에 내홍 겪는 민주당	100
4	5	"반도체 살릴 특별법 서둘러야" 업계-학계 호소	100
...	...	...	...
1454	1898	이스라엘-팔 총돌 사망자 90명까지...유대인-아랍인 집단 폭력(종합)	500
1455	1899	싱가포르도 코로나 경고음...지난해 7월 이후 지역감염 최다	500
1456	1900	美서 안센 백신 맞은뒤 혈전 부작용 28명으로 늘어...남성도 6명	500
1457	1901	18개 금빛 거울 펼쳐졌다...우주의 눈 제임스 웹 망원경 곧 발사	600
1458	1902	"아이폰13 두께 더 두꺼워진다"...카메라 기능 대폭 강화	600
...	...	...	...
0	1	국민의힘 주호영 당 대표 권한대행 겸 원내대표(왼쪽)와 안철수 국민의...	
1	2	/자료제공 = NBS[서울경제] 차기 대권후보 적합도 조사에서 이재명...	
2	3	심상정 정의당 의원이 29일 오전 서울 여의도 국회 소회의실에서 토지초...	
3	4	더불어민주당이 당내 일부 강성 지지자들이 보내는 문자 폭탄...	
4	5	"반도체 대기업 혜택 법안이라고 색안경 끼고 보지 말고, 국민 삶의 ...	
...	...	...	...
1454	1898	네타냐후 "무정부 상태...사태 진정 위해 군 병력 투입"[AP=연합뉴...	
1455	1899	26일 홍콩과 트래블 버블·6월초 상그릴라 대화 앞두고 집중[EPA...]	
1456	1900	3명 숨지고 4명 입원 중...통계상 30~49세 여성에서 혈전 위험성...	
1457	1901	[사이언스샷] 미국 노스롭 그루먼사의 시험장에서 진행된 제임스 웹 우주...	
1458	1902	두께 호불호 갈릴 수 있다는 분석도 [사진=맥루머스] 올 하반기 출시할 ...	



count	title	category	#
0	10	"뚝뚝, 저녁거리 왔어요"...CJ대한통운 식품 택배 전년比 50%↑	101
1	11	공시대상기업 7개 증가...쿠광 김법석 총수 지정 피해	101
2	12	간선급행버스, 세종~공주·청주 신설...비수도권 광역교통 확충	101
3	13	출서기·싸늘이 다 막는다...e-프리퀀시 시작하는 스타벅스의 각오	101
4	14	홍남기 "2025년까지 100대 핵심서비스 표준 개발"	101
...	...	...	...
291	1879	전세계 코인 투자자 우뚝한 머스크..."증시였다면 감옥행"	300
292	1880	[단독]현대차, 투싼 생산 중단한다...車 반도체 보릿고개 본격화	300
293	1881	"그때 샀어야..." 2년 전 매매가 뛰어넘는 전셋집 쏟아졌다	300
294	1882	국민 10명 중 8명 "무주택·서민 실수요자 LTV 더 우대해야"	300
295	1883	LH 수사 뛰어난 검찰...전현직 직원 유착 의혹 정조준	300
...	...	...	...
0	1	◎ 뉴스1(서울=뉴스1) 윤다정 기자 = 지난해 CJ대한통운을 통해 ...	
1	2	[CBS노컷뉴스 지영한 기자]공정거래위원회, 연합뉴스시중유동성 증가와...	
2	3	부산 미술~가락, 대구 안심~경산 임당 광역도로 등도 신설부산·울산권...	
3	4	[디지털타임스 김아름 기자] 스타벅스가 오는 5월부터 여름 e-...	
4	5	홍남기 국무총리 직무대행이 29일 서울 여의도 수출입은행에서 열린 제...	
...	...	...	...
291	1879	머스크의 배신...비트코인 폭락wr "테슬라, 비트코인 안받겠다" 돌발 트윗...	
292	1880	6일 오전 현대자동차 울산공장 명촌정문에서 남품 차량들이 드나들고 있다. 현...	
293	1881	시흥의 한 아파트단지. 한경DB 전세 공급난이 장기간 지속되면서 전세가격이 2...	
294	1882	서울의 한 은행 대출창구. 뉴스1 DB ◎ News1 (서울=뉴스1) 박...	
295	1883	LH 출신, 신도시 개발 개입 정황 포착...강제수사 돌입[연합뉴스TV...]	

[296 rows x 4 columns]

목적 기사와 같은 카테고리 기사를 대상으로 축소



## 기사 추천 실행 과정 및 결과 1

코사인 유사도 계산:  
(1일 경우 동일)

```
[[1.          0.32268637 0.11201012 ... 0.0565903  0.46978497 0.13400094]
 [0.32268637 1.          0.27572476 ... 0.0815841  0.24604124 0.11169309]
 [0.11201012 0.27572476 1.          ... 0.06420185 0.09525956 0.33321405]
 ...
 [0.0565903  0.0815841  0.06420185 ... 1.          0.07851298 0.21089824]
 [0.46978497 0.24604124 0.09525956 ... 0.07851298 1.          0.10233041]
 [0.13400094 0.11169309 0.33321405 ... 0.21089824 0.10233041 1.          ]]
(296, 296)
```

추천 기사 결과:

```
<< 유사도 상위 5개 >>
[(200, 0.6631317464437189), (217, 0.48130493360243387), (254, 0.41655099180793015), (245, 0.404905218223
2526), (264, 0.3921073428055738)]

<< 제목이 "아파트 공시가격 기초자료 나왔지만...여전히 부족한 설명"인 기사와 유사한 5개 기사 >>
200      새 임대차법 9개월...전세 줄고 월세·반전세 늘었다
217      한섬, 영업이익 54% 급증 어닝 서프라이즈
254      아파트 전셋값이 71억원...3.3㎡당 1억원 넘는 역대 최고액
245      버티기 들어간 다주택자...아파트 매물 줄고 거래도 절벽
264      "서울 집값 더 오른다는데 누가 팔겠나"...집주인들 배짱
Name: title, dtype: object
```

추천 기사의 db내  
count값리턴:

```
Out[31]: [1375, 1920, 1469, 1688, 1630]
```



## 기사 추천 실행 결과 2

추천 목적 기사:

count	8
title	필수부대 조종사 · 함정도 첫 백신접종..."안심하고 임무 매진"
category	100
content	30세 이상 장병 AZ접종 일제 시행...전투복 차림 · 태극기 향해 경례 ...

추천 가능 기사:  
(dataframe)

	count	title	category	#
0	1	주호영 · 안철수 "합당 공감...큰 어려움 없을 것"		100
1	2	대권후보 적합도 이재명 24%vs윤석열 23%...이낙연은 또 한자릿수		100
2	3	'초과이득 절반 세금으로' ...심상정, 폐지 23년 만에 토초세법 발의		100
3	4	문자 폭탄이 뭐길래... 강경파에 내홍 겪는 민주당		100
4	5	"반도체 살릴 특별법 서둘러야" 업계-학계 호소		100
...	...	...	...	...
1468	1913	유력 대권주자 가상 양자대결, 이재명 42% vs 윤석열 35.1% [한국갤럽]		100
1502	1947	문대통령 "당이 주도적으로 정책 마련하는 게 필요"(종합)		200
1503	1948	당청 先 후보사퇴 입모았다 대통령 말에 원점...막후 들어보니		200
1504	1950	백신 · 안보 · 경제...한미정상회담, 한미관계 최대 변곡점		200
1505	1951	집 신발장에 숨겨진 노트북...이상직 구속의 결정타였다		200
...	...	...	...	...
0		국민의힘 주호영 당 대표 권한대행 겸 원내대표(왼쪽)와 안철수 국민의...		
1		/자료제공 = NBS[서울경제] 차기 대권후보 적합도 조사에서 이재명...		
2		심상정 정의당 의원이 29일 오전 서울 여의도 국회 소통관에서 토지초...		
3		더불어민주당이 당내 일부 강성 지지자들이 보내는 문자 폭탄...		
4		"반도체 대기업 혜택 법안이라고 색안경 끼고 보지 말고, 국민 삶의 ...		
...		...		
1468		여야 유력 대권주자 간 가상 양자대결에서 이재명 ...		
1502		興 지도부와 간담회..."당청 원팀으로 남은 1년 성공적 마무리""부동...		
1503		"민주 지도부, 주말 당정청 협의회사 일부 인사 낙마론 제기"(서울...		
1504		조 바이든 미국 대통령, 문재인 대통령.© News1 최수아 디자이너 (...		
1505		이스타항공 창업주로서 횡령 · 배임 혐의를 받는 무소속 이상직 의원이 27일 전북...		

[334 rows x 4 columns]

목적 기사와 같은 카테고리 기사를 대상으로 축소



## 기사 추천 실행 결과 2

### 추천 기사 결과:

<< 유사도 상위 5개 >>

```
[ (37, 0.8202125030555731), (167, 0.7472717045816623), (104, 0.7034153186018653), (45, 0.657987654250131), (68, 0.6185512789072476) ]
```

<< 제목이 "필수부대 조종사·함정도 첫 백신접종..."안심하고 임무 매진""인 기사와 유사한 5개 기사 >>

```
37      군인·군무원 2만명, 아스트라 접종 거부
167     상반기 1천300만명 이상 접종 가능하려면...정부가 밝힌 전략은
104     문 대통령, 11월 집단면역 달성 자신감
45      文, AZ 접종 간격 12주인데 5주 만에 2차 맞은 이유는?
68      문 대통령, 아스트라제네카 백신 2차 접종
```

### 추천 기사의 db내 count값리턴:

```
Out[30]: [140, 1222, 794, 207, 411]
```



# Part 6.

---

## 활용방안 및 기대효과





## 활용 방안 및 기대효과

바쁜 직장인들이 빠르게 기사를  
읽을 수 있어서 시간이 절약될 것  
이다.

카드뉴스로 기사를 제공하여  
가독성을 향상시킨다.

주목 받는 이슈에 대한 기사를  
주변인들에게 설명하고 싶을  
때 간편하고 유용하게 사용할  
수 있다.

추천시스템을 통해 사용자가  
원하는 기사를 찾는 소요시간  
을 줄일 수 있다.



# Part 7.

---

시연 영상







감사합니다

