

# 2021 한이음 ICT멘토링

프로젝트명

빅데이터 기반 주식 및 뉴스 분석 시스템 구축

# 요 약 본

프로젝트 정보	
주제영역	<input type="checkbox"/> 생활 <input type="checkbox"/> 업무 <input type="checkbox"/> 공공/교통 <input checked="" type="checkbox"/> 금융/핀테크 <input type="checkbox"/> 의료 <input type="checkbox"/> 교육 <input type="checkbox"/> 유통/쇼핑 <input type="checkbox"/> 엔터테인먼트
기술분야	<input type="checkbox"/> IoT <input type="checkbox"/> 모바일 <input type="checkbox"/> 데스크톱 SW <input checked="" type="checkbox"/> 인공지능 <input type="checkbox"/> 보안 <input type="checkbox"/> 가상현실 <input checked="" type="checkbox"/> 빅데이터 <input type="checkbox"/> 자동제어기술 <input type="checkbox"/> 블록체인 <input type="checkbox"/> 영상처리
달성성과	<input checked="" type="checkbox"/> 논문게재 및 포스터발표 <input type="checkbox"/> 앱등록 <input type="checkbox"/> 프로그램등록 <input type="checkbox"/> 특허 <input type="checkbox"/> 기술이전 <input type="checkbox"/> 실용화 <input checked="" type="checkbox"/> 공모전( 한이음 ICT 멘토링 공모전 ) <input type="checkbox"/> 기타( )
프로젝트명	빅데이터 기반 주식 및 뉴스 분석 시스템 구축
프로젝트 소개	빅데이터를 활용해 주식 현황 정보 및 주가 예측, 뉴스 감성분석 서비스를 제공하는 프로젝트
개발배경 및 필요성	<ul style="list-style-type: none"> <li>• 텍스트 자료를 이용하여 경제 문제를 분석한 해외의 연구들은 웹 검색어 빈도분석, 소셜미디어 텍스트 분석, 뉴스기사분석 등 다양한 방법을 활용 중에 있음.</li> <li>• <b>한국인의 금융이해력은 OECD 평균 이하</b>. 다양한 소득층의 금융 이해력을 높이기 위해 텍스트 자료를 이용하여 주식 뉴스 분석 시스템을 제공할 필요성이 있음.</li> <li>• 투자자들이 특정 커뮤니티 글이나 종목 토론방을 보고 <b>감정적으로 투자 결정을 내리는 경향이 강해지면 주가 변동 위험성이 커지기 때문에 합리적인 의사결정을 할 수 있도록 돕는 빅데이터 시스템이 필요함.</b></li> </ul>
프로젝트 주요기능	① 다양한 연령층이 합리적인 의사결정을 할 수 있도록 <b>주식 뉴스 분석 시스템</b> 제공 ② 기업과 관련된 <b>경제 뉴스 텍스트 분석</b> ③ 빅데이터 트렌드를 제공하고 감성분석을 이용해 <b>종목별 주가를 예측</b> ④ <b>주가 전망을 예측</b>
작품의 기대효과 및 활용분야	<b>주가 지수 예측</b> - 뉴스 감성 분석을 추가하여 정성적 측면의 분석 <b>지능형 투자 결정</b> - AI 인공지능을 이용한 지능형 투자 자문 및 투자 결정 <b>충동적인 투자 위험성 완화</b>

# (본문) 프로젝트 결과보고서

## I. 프로젝트 개요

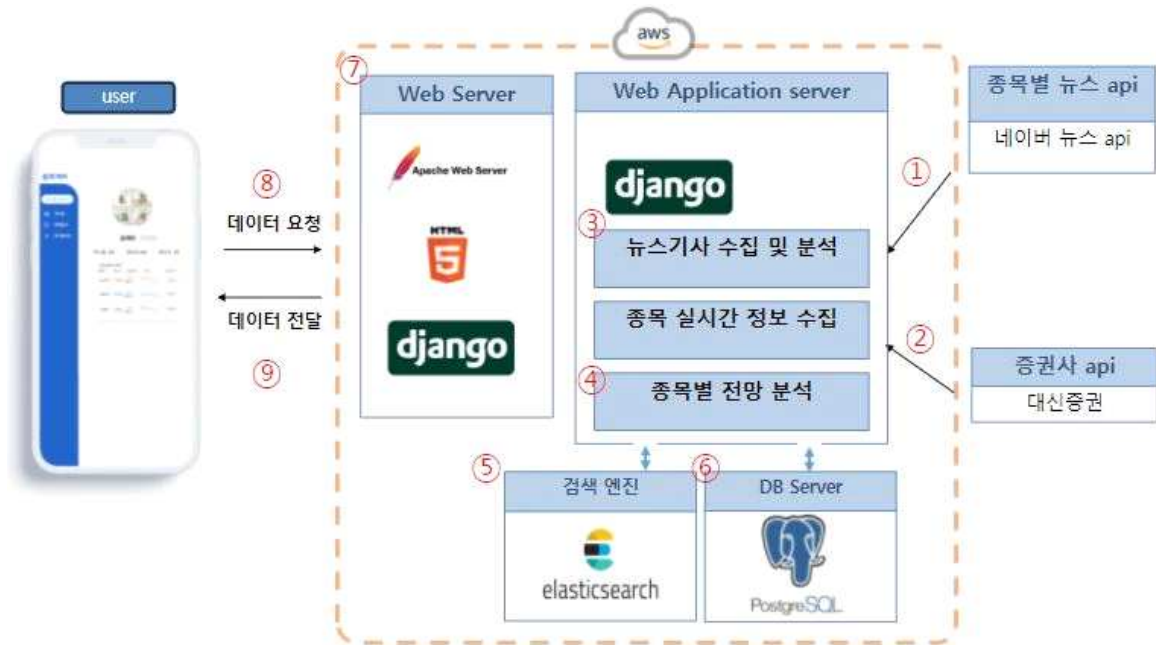
### 가. 프로젝트 소개

- 빅데이터를 활용하여 경제 및 주식 현황 정보 제공 서비스를 구축하는 프로젝트
- 뉴스를 분석하여 뉴스에 나온 키워드가 주가에 미친 영향 학습 및 예측
- 딥러닝을 활용해 특성을 생성하고, 주가 전망을 예측에 활용

### 나. 개발배경 및 필요성

- 경제 정책 분석에 빅데이터를 활용하는 방안에 대해 최근 많은 논의가 이루어지고 있음
- 텍스트 자료를 이용하여 경제 문제를 분석한 해외의 연구들은 웹 검색어 빈도 분석, 소셜미디어 텍스트분석, 뉴스 기사 분석 등 다양한 방법을 활용 중에 있음.
- 2018년 우리나라 성인 2400명을 대상으로 조사한 한국인의 금융 이해력은 62.2점으로 OECD평균 (2015년 64.9점) 보다 낮은 수준이다. 특히 금융이해력은 소득별 차이가 컸으며, 소득이 낮을 수록 금융이해력이 떨어졌다. 따라서 다양한 연령층과 소득층들이 올바른 금융가치관을 형성 하고 합리적인 의사결정을 할 수 있도록 주식 뉴스 분석 시스템을 제공할 필요성이 있음.
- 해외 주식 정보에 대한 접근성 완화  
최근 2030 세대의 해외주식 투자 성향이 급격하게 증가했다. 이러한 경향으로 이른 바 '테슬라 쓸림현상'이 발생해 전반적으로 위험투자 성향이 높아지고 있다. 더 많은 투자 기회를 제공한다는 점에서 해외주식시장 확대는 필요하지만 국내 주식 시장 정보에 비해 정보 접근성이 낮다. 자칫하면 주가변동 위험에 크게 노출될 수 있기 때문이다. 이러한 상황에서 가장 필요한 것이 빅데이터 시스템이다. 빅데이터 주식 정보 시스템을 구축한다면 해외주식 투자 정보에 대한 접근성을 완화하는 데 큰 기여를 할 수 있다.

## 다. 작품 구성도



## 라. 작품의 특징 및 장점

- 빅데이터를 활용해 주식을 예측하는 데 가장 큰 장점이 있음. 주가를 예측함으로써 이익을 최대화하고 손실을 최소화함.
- 감성 분석을 활용하여 주식 종목에 대한 긍정, 부정 의견을 파악하고 최근 동향을 살피는 데 도움을 줌.

## II. 프로젝트 수행결과

### 가. 주요기능

구분	기능	설명
S/W	경제 뉴스 분석 및 추천 시스템	AI 머신러닝을 통한 주요 경제뉴스 추천 및 기업 분석
	주가 예측 시스템	빅데이터 트렌드 및 감성분석을 이용한 종목별 주가 예측
	웹페이지	프로젝트 결과물을 웹페이지로 구현 / 실행 / 테스트

### 나. 프로젝트 개발환경

구분	항목	적용내역
S/W 개발환경	OS	Window 10, 64bit(x64 기반 프로세서) Mac Os, Apple M1 Ubuntu Server 18.04 (Linux)
	개발환경(IDE)	PyCharm
	개발도구	Git, GitHub, GitLab
	개발언어	Python, HTML, CSS, JavaScript frontend는 HTML, CSS, JS 등으로 구현, backend는 Python 이용, 알고리즘 분석은 JupyterNotebook 이용.

### 다. 장비(기자재/재료) 활용

번호	품명	작품에서의 주요기능
1	해당사항 없음	-
2	해당사항 없음	-
3	해당사항 없음	-
4	해당사항 없음	-

### 라. 프로그램 작동 동영상

<https://youtu.be/DE5gPhmNIAs>

## 마. 결과물 상세 이미지

동학개미

종목분석

매매일지

포트폴리오

관심종목

감성분석

종목명	종목코드	가격	종목률	상승/하락 예측	확률
유한양행	A000100	59200	-0.01%	상승	80%
SK하이닉스	A000660	92700	-0.01%	하락	77%
신세계	A004170	254000	-0.01%	상승	86%
현대차	A005380	210500	0.03%	상승	59%
POSCO	A005490	317000	-0.01%	하락	22%
삼성전자	A005930	69300	-0.03%	하락	94%
한화솔루션	A009830	44750	0.06%	하락	87%
HMM	A011200	29300	0.01%	상승	66%
현대모비스	A012330	270000	0.01%	하락	93%
아시아나항공	A020560	24150	-0.01%	하락	89%





## 바. 달성성과

<input checked="" type="checkbox"/> 논문게재 및 포스터발표	게재(발표)자명	논문(포스터)명	게재(발표)처	게재(발표)일자
	김지현	설명 가능 인공지능 기법을 활용한 주가 전망 예측	한국정보처리학회	2021. 11. 3.
<input type="checkbox"/> 앱(APP) 등록	등록자명	앱(APP)명	등록처	등록일자
				2017. 00. 00.
<input type="checkbox"/> 프로그램 등록	등록자명	프로그램명	등록처	등록일자
				2017. 00. 00.
<input type="checkbox"/> 특허/실용신안 출원	출원자명	특허/실용신안명	출원번호	출원일자
				2017. 00. 00.
<input type="checkbox"/> 기술이전	기술이전기업명	기술명	금액	이전일자
				2017. 00. 00.
<input checked="" type="checkbox"/> 공모전	구분(교내/대외)	공모전명	수상여부(출품/수상)	상격
	대외	한이음 공모전	1차 합격	x
<input type="checkbox"/> 실용화	#실용화한 내용에 대한 구체적 작품설명			
<input type="checkbox"/> 기타				

### Ⅲ. 프로젝트 수행방법

#### 가. 업무분장

번호	성명	역할	담당업무
1	임진섭	멘 토	- 프로젝트 진행 관련 멘토링, 개발 톨과 방향 제안
2	정수민	팀 장	- 프로젝트 관리, 웹 프로그래밍, 증권사 API를 사용하여 일일 주가 데이터 조회 서비스 구현
3	조설아	팀 원2	- 주가 예측 알고리즘 실험, 머신러닝 모델 생성 및 저장, 뉴스 데이터 크롤링 및 저장
4	이연수	팀 원3	- Django를 사용한 웹사이트 구현, 주가예측 알고리즘 실험
5	김지현	팀 원4	- 주가 예측 알고리즘 실험, 뉴스 데이터 크롤링 및 저장, 감성 분석, 자연어 처리, mongo db연결
6	안정은	팀 원5	- 주가예측 알고리즘 실험, 웹 프로그래밍

#### 나. 프로젝트 수행일정

구분	추진내용	수행일정									
		3월	4월	5월	6월	7월	8월	9월	10월	11월	
계획	프로젝트 팀 구성 및 주제 선정 서비스 기획 프로젝트 일정 계획										
분석	선행 연구 분석 제공할 서비스 분석										
설계	웹 UI/UX 설계										
	데이터 수집 및 주가예측 시스템										
개발	웹 구현										
	뉴스 분석 시스템										
	주가 예측 시스템										
테스트	테스트 및 보완										
종료	결과 보고서 및 공모전 출품										
오프라인 미팅	오프라인 멘토링										



#### 다. 문제점 및 해결방안

##### ○ 프로젝트 관리 측면

문제점	해결방안
시간 조율 문제	⇒ 공유 문서를 이용해 일주일 단위로 일정을 조율하여 미팅 시간을 정하고 세부계획을 세움 ⇒ 방학에는 주 4회이상 온라인 미팅
코로나로 인한 오프라인 미팅의 어려움	⇒ Zoom, Google Meet, Google Drive, Micro Teams, Kakao Talk, gather town 등 온라인 플랫폼 적극 활용
깃 커밋 충돌 등 사용방법이 어려움	⇒ 팀원들과 관련 자료를 공유하고, 한이음에서 제공한 강의를 공부하여 해결
모델 배포 어려움	⇒ Django rest framework 공부, 멘토링을 통하여 해결

##### ○ 작품 개발 측면

문제점	해결방안
금융 뉴스 데이터 과다 축적시 데이터베이스의 과부하가 생길 수 있음	⇒ 2011-2020년으로 기간 제한
새로 다루는 기술과 프로그램에 대한 학습을 요함으로 인해 프로젝트 지연	⇒ 스파르타 코딩클럽 강의 수강, 매주 서적 및 논문 스터디로 진행 차질 최소화 ⇒ 매 주 정해진 범위를 정하여 프로젝트에 필요한 기술을 학습 ⇒ 온라인 미팅으로 학습한 내용을 이해했는지 피드백하는 시간을 가짐
autoencoder 압축 변수 선택	⇒ 멘토님께 가격데이터보다는 기술지표를 압축변수로 설정하는 것이 예측 정확도를 높일 수 있다는 피드백을 받음
모델 배포 방법이 어려움	⇒ 멘토링으로 해결
각자의 로컬 환경에서 개발하는 상황	⇒ 각자의 PC에서 작업하다보니 공동 개발 환경을 구축할 필요성을 인식. 멘토링과 AWS강좌의 도움을 받아 AWS EC2로 해결방안을 강구함.

## IV. 기대효과 및 활용분야

- 주가 중장기 전망 예측
  - 단순히 주가의 다음 날 가격이 아닌 주가의 중장기적인 전망을 예측
- 주가 정보 제공
  - 차트뿐만 아니라 뉴스, 감성을 분석하여 정성적 측면의 분석
- 지능형 투자 결정
  - AI 인공지능을 이용한 지능형 투자 자문 및 투자 결정
- 충동적인 투자 위험성 완화

## V. 참고자료

### 가. 참고 및 인용자료

- 임청수, "A study on Stock Trend Determination in StockTrend Prediction", 한국컴퓨터정보학회논문지, Vol.25, No. 12, pp.35-44, Dec.2020.