

자연어 처리 기반의
투자 분석 및 예측 시스템 개발

TURN-AROUND



INDEX



1. 프로젝트 배경 및 필요성



산학 프로젝트 개요



과제명

자연어 처리 기반의 투자 분석 및 예측 시스템 개발



팀명

TURN-AROUND



멘토

정좌연 PE



멘티

이지훈, 강민재, 이문형, 김서정, 구병진



과제 수행기간

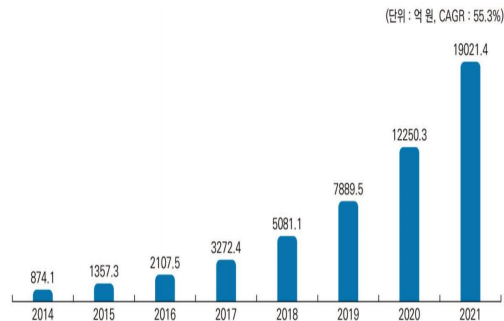
2020.08.03 ~ 2020.11.06

1. 프로젝트 배경 및 필요성

기존 로보어드바이저의 한계 극복

- 자연어 처리와 관련한 한계점 존재
- 전문화를 통한 세밀한 모델 적용 요구

로보 어드바이저 산업의 활황



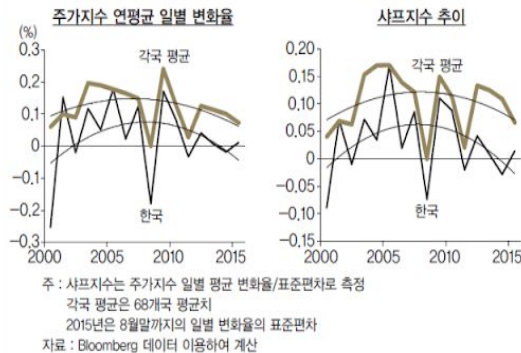
자료 : 김광훈, 인공지능 자산운용시스템, 한국과학기술정보연구원, KISTI Market Report, 2016.

1. 프로젝트 배경 및 필요성

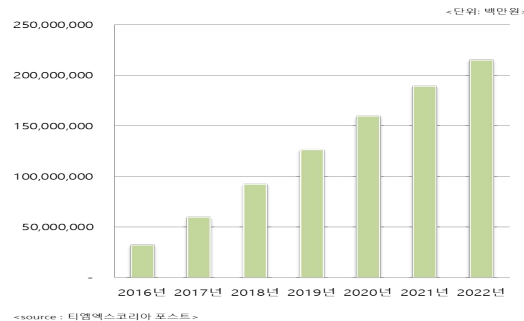
연관산업의 추세를 반영한 정밀성 필요

투자 패러다임 변화 : 비대면 투자 성장

- 금융 시장의 전염효과
- 기술적 분석 보완
- 자본시장법 시행령, 금융투자업 규정 개정안
- 투자비용 감소 가능



P2P 핀테크 시장규모



1. 프로젝트 배경 및 필요성

자산운용과 투자의 진입장벽 해소

언어 감성분석을 통한 주가 선제적 예측

- 젊은층 및 노년층에 겨냥
- 주식의 매수 및 매도 시점에 대한 혜안
- 언어 분석을 통해 민감한 반응
- 주식 매매에 대한 리스크 최소화



2. 프로젝트 목표 및 내용

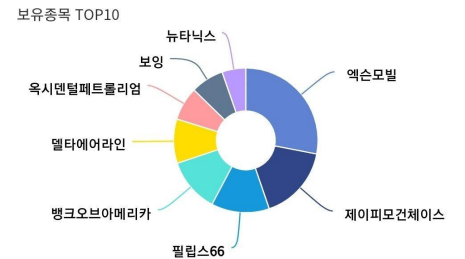


2. 프로젝트 목표 및 내용

로보어드바이저 ROBO-ADVISOR

빅데이터와 알고리즘에 기반한 컴퓨터의 통계적 분석과 전략으로
투자자의 투자성향과 목표에 맞춰 개인화된 포트폴리오를 제시하고
시스템적으로 자산을 관리

주식, 경제, 기업 데이터를 활용한 AI-ADVICE SYSTEM



2020.04.09 기준

종목명	수량	평가금액	보유비중
엑슨모빌	400	17,252.00	28.05%
제이피모건체이스	100	10,276.00	16.71%
필립스66	130	7,997.60	13.00%
뱅크오브아메리카	300	7,458.00	12.12%
델타에어라인	250	6,097.50	9.91%
옥시덴탈페트롤리엄	300	4,608.00	7.49%

2. 프로젝트 목표 및 내용

로보어드바이저 : ROBOT(로봇)과 ADVISOR(자문가)의 합성어

빅데이터와 알고리즘에 기반한 컴퓨터의 통계적 분석과 전략으로
투자자의 투자성향과 목표에 맞춰 개인화된 포트폴리오를 제시하고
시스템적으로 자산을 관리



3. 추진 전략 및 방법



3. 추진 전략 및 방법

1단계 주식관련 데이터 분석

뉴스 및 리포트 활용	시계열 데이터 분석	기업 관련 데이터
종목별 주가 분석 보고서, 뉴스 등 텍스트 데이터 크롤링	글로벌 지표 (S&P 500, KOSDAQ, KOSPI) 수집	산업, 섹터 분석 (경제연구소자료, 산업별 통계자료)
투자 키워드 텍스트 마이닝	API를 활용한 증권사, KRX 데이터 크롤링	내부 분석 (기업 공시 자료, 신용평가 자료 활용)
투자 키워드 군집화 및 탐색적 분석	적정 주가 및 추세이탈 징후 파악	국내외 특허 / 기술 / 연구 조사 보고서

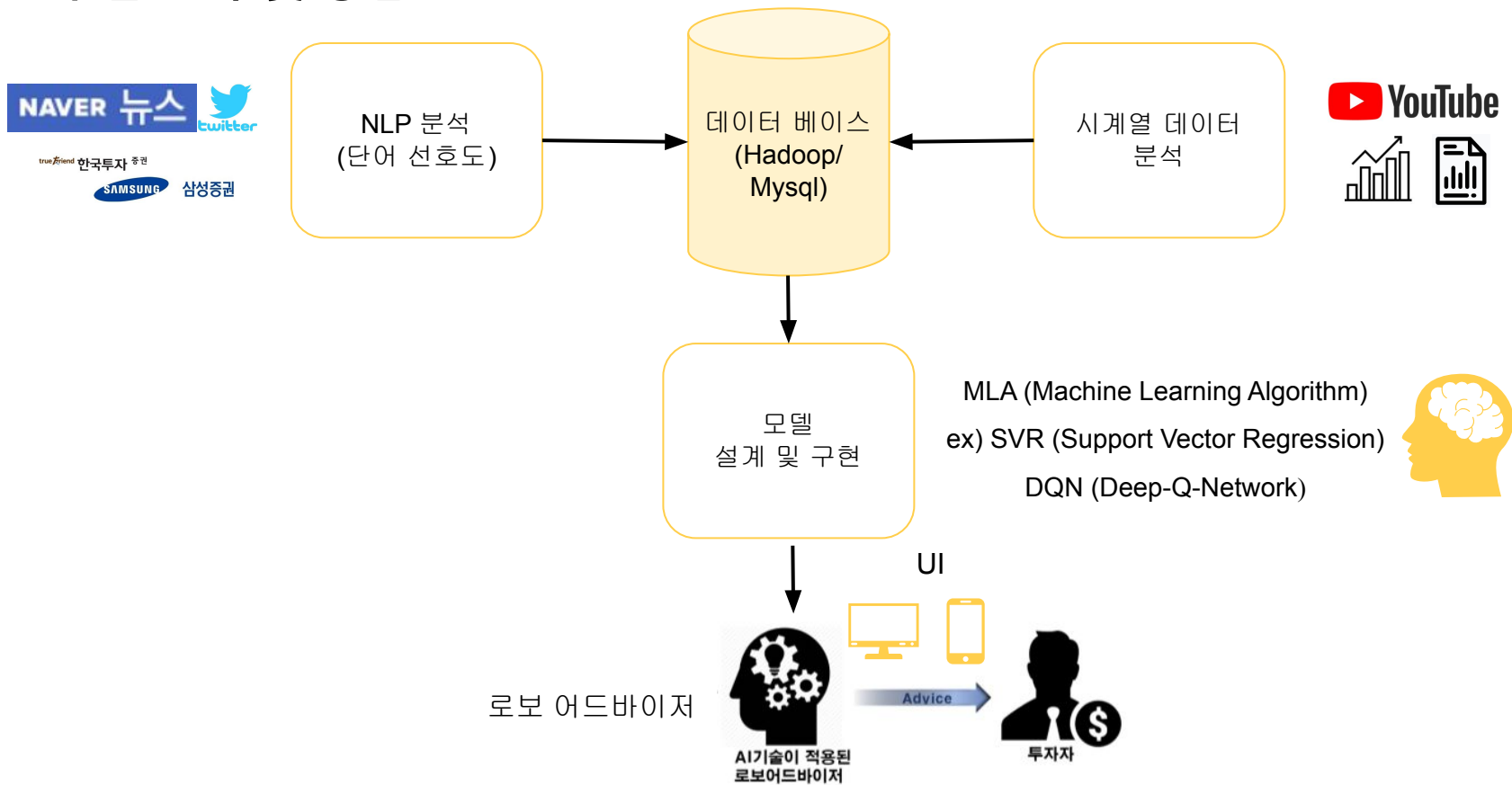
2단계 예측모델 설계

데이터 수집 및 전처리	알고리즘 설계	데이터베이스 및 UI 설계
로그 수집 분석	마르코프 결정 프로세스	분석 시각화 구성
데이터 정제 처리	DQN을 활용한 주가예측 알고리즘 설계	ER-diagram / 테이블 명세서
훈련, 검증, 평가 데이터 셋	성능 평가 지표	DB 구성도 / 화면 구성도

3단계 서비스 구현

알고리즘 구현	테스트
자연어 처리 엔진 구현	알고리즘 실험 및 하이퍼 파라미터 튜닝
상승/하락, 적정주가, 매수매도 포인트 예측	로보어드바이저 프로토타입 개발

3. 추진 전략 및 방법



4. 추진 일정 및 성과물



4. 추진 일정 및 성과물

프로젝트 수행단계		수행내용	산출물	
[1단계] 계획 및 분석	도메인(금융업) 관련 지식 습득	- 투자지표, 재무제표, 차트분석, 경제용어 공부	-	- 착수보고서
	선행연구 및 기술동향 조사 분석	- 논문 조사 (자연어 처리 및 머신러닝, 딥러닝 기반의 주가 예측) - 로보어드바이저 기술 동향 분석	- 국내외 특허 / 기술 / 연구 조사 보고서	
	종목별 (산업별) 주가 분석	- 기업 분석 - 데이터 수집 (투자지표 및 관련있는 문서 크롤링) - 적정 주가 및 추세 이탈 징후 파악	- 종목별 주가 분석 보고서 - 원시 데이터	
[2단계] 설계	데이터 수집 및 전처리	- 로그 수집 분석 - 데이터 정제 처리	- 훈련, 검증, 평가 데이터 셋	- 중간보고서 - 투자분석 시스템 구성도
	알고리즘 설계	- 투자 키워드 텍스트 마이닝 설계 - 투자 키워드 군집화 및 탐색적 분석 - 자연어 처리 및 머신러닝 (딥러닝) 기반의 주가 예측 알고리즘 설계	- 알고리즘 설계 구성도 - 성능 평가 지표 - 분석 시각화 구상	
	데이터베이스 및 UI 설계	- 데이터베이스 설계 - 화면 설계	- ER-diagram / 테이블 명세서 - DB 구성도 / 화면 구성도	
[3단계] 구현	알고리즘 구현	- 알고리즘 구현	- 워드 클라우드 / 투자 키워드 트렌드 - 주가 예측 / 적정 주가 산출 / 추세 이탈 징후 파악	- 최종보고서 - 로보어드바이저
	데이터베이스 및 UI 구현	- 데이터베이스 및 UI 구현	- 로보어드바이저 프로토타입	
	테스트	- 모델 최적화 및 평가 (알고리즘 실험 및 하이퍼 파라미터 튜닝)	- 로보어드바이저	
	최종보고서 작성	- 최종보고서 작성	- 최종보고서	

4. 추진 일정 및 성과물

수행내용		수행일정												
		8월				9월				10월				11월
		1	2	3	4	1	2	3	4	1	2	3	4	1
보고	1. 착수보고회 / 중간보고회 / 최종보고회 / 정기 미팅 (매주 수요일 18:00 ~ 22:00)			14				25						06
[1단계] 계획 및 분석	2. 도메인 관련 지식 습득													
	3. 국내외 특허 / 기술 / 연구 조사 보고서													
	4. 종목별 주가 분석 보고서													
[2단계] 설계	5. 데이터 수집 및 전처리													
	6. 투자 키워드 텍스트 마이닝 설계													
	7. 투자 키워드 군집화 및 탐색적 분석													
	8. NLP 및 머신러닝,딥러닝 기반의 주가 예측 알고리즘 설계													
	9. 데이터베이스 및 UI 설계													
[3단계] 구현	11. 알고리즘 구현													
	12. 데이터베이스 및 UI 구현 (로보어드바이저 프로토타입)													
	13. 알고리즘 최적화 및 평가													
	14. 로보어드바이저 개발 완료													

“참고문헌”

『국내외 로보어드바이저 동향 및 현황 분석』 - 이근영 , 금융보안원
보안연구부

『금융시장의 충격에 대한 서베이』 - 오동철, 임현준 , 한국은행 경제연구원

『A Discrete Stock Price Prediction Engine Based on Financial News』
- Rob Schumaker University of Texas at Tyler

спасибо 谢谢
GRACIAS

THANK YOU

ありがとうございました MERCI

DANKE धन्यवाद

شُكراً OBRIGADO