

3주차 과제

제출기한: 2020년 4월 9일 (목) 23:59:59

팀 번호: 9

1. 프로젝트 주제

팀원들의 2주차 과제를 바탕으로, 팀의 최종 프로젝트 주제를 선정한다. 진행되는 프로젝트는 Github에 지속적으로 코드를 업로드하여 지도교수 및 조교가 진행상황 및 팀원의 기여도를 확인할 수 있도록 한다. 따라서 모든 팀원이 Github 계정을 만들고 팀원 중 한 명(팀장)이 저장소(repository)를 프로젝트 이름으로 만든 뒤, collaborator로 나머지 팀원을 추가하도록 한다. (파일 마지막 장 참고자료 참고) 또는 프로젝트 이름의 계정을 만들어서 그 안에 저장소를 만들고, 해당 저장소에 모든 팀원을 collaborator로 추가해둬도 됨.

프로젝트 주제: 딥러닝을 이용한 주식가격 예측 웹서비스

프로젝트 Github 저장소 주소: <https://github.com/cjssla1/projectJucker>

팀원1(팀장)	이름: 김경중	Github ID: cjssla1
팀원2	이름: 김현재	Github ID: KHJae
팀원3	이름: 김진용	Github ID: w987987
팀원4	이름:	Github ID:

2. 프로젝트 계획

확정된 프로젝트 주제를 구현하기 위한 계획을 작성한다. 아래 항목은 글씨 크기 9인 자유 포맷으로 작성 하되, 분량은 전체 2장을 넘어야 함. 2주차 과제에서 작성한 내용을 활용 가능함.

- **필요성** : 해당 프로젝트를 활용할 수 있는 상황 및 대상 등 실제 사례를 바탕으로 한 필요성 서술.
- **독창성** : 비슷한 주제로 기존에 개발된 프로그램 등을 조사하여 작성하고, 본 프로젝트에서 개발할 내용은 어떤 점이 다른지 서술.
- **필요 기술** : (1) 팀원들이 사용 가능한 프로그래밍 언어에 기반하여, 본 프로젝트에 사용할 프로그래밍 언어를 1개 이상 작성. (2) 프로젝트에 활용 가능한 오픈소스 라이브러리를 조사하고 본 프로젝트의 어떤 부분에 사용할 수 있는지에 대한 설명과 출처(링크 등)를 작성.
- **월별 계획 설정**: 4월~10월에 대해 월별 계획을 기술. 참고: 6월 중순에는 이번 학기 중간 결과 발표 예정이며, 최근 4년간 캡스톤디자인 전시회는 11월 7-14일 경 진행되었음(올해 일정은 미정). 따라서 최종적으로 10월까지의 마무리되는 것을 목표로 할 것.
- **주차별 계획 설정**: 남은 학기 (4~6월)에 대해 주차별 계획을 기술하고, 각 팀원의 예상 역할을 나눠볼 것. 추후 팀원 간 평가도 성적에 반영될 예정이니 각자 작업 분배와 참여를 확실히 할 것.

1) 프로젝트 내용

딥러닝을 이용한 주식가격 예측 웹서비스

2) 필요성

메리츠 자산운용의 대표 존리는 "주식에 투자한다는 것은 내가 투자한 회사의 주인이 된다는 것, 회사 임직원이 나를 위해 일하게 한다는 것, 회사가 창출하는 부를 분배받을 권리를 갖는다는 의미를 지닌다"며 "창업 외에 자본가가 되는 유일한 방법이 주식 투자"라고 말했다. 한 번에 많은 금액을 투자할 필요는 없다고 덧붙였다. 물가상승률에 비해 금리상승률은 크지 않습니다. 그 말은 단순 저축은 오히려 시간이 갈수록 돈의 가치가 떨어진다는 뜻입니다. 예전엔 천원은 버스를 탈 수 있었으나 지금 천원으론 버스를 탈 수 없습니다. 즉 일반 근로자가 중산층이 되기 위해서 주식투자는 불가피합니다. 그러나 아무런 준비도 없이 직관만을 의존해서 주식시장에 단타를 노리는 것은 무모한 짓입니다.

18년 7월 30일 기사 "개미 사들인 종목 '쭈박'... 외인,기관 떠나는데 개미만 '피눈물', ", 17년 1월 2일 기사 "코스피 개인 평균 수익률 -33.07%, 기관투자자 수익률 15.69%..." 이런 기사들을 통해 개인 투자자들은 성공하기 쉽지 않다는 것을 알 수 있습니다. 이유가 무엇일까요? 사실 증권기관이나 개인이나 주식시장 정보에서 얻을 수 있는 데이터는 아주 많은 차이가 있지 않습니다. 그런데도 차이가 심한 이유는 무엇일까요? 바로 시선의 차이입니다. 기관은 단순 현재 데이터만 보지 않고 이전의 비슷한 데이터를 참고하며 또한 한 사람이 사고 팔고를 결정하지 않고 여러 사람이 여러 시선으로 보고 사고 팔고를 결정합니다. 특정주를 전임한다고 해도 단순 개인 직감이 아닌 여러 데이터를 참고하여 결정합니다. 그에 비해 개인투자자는 대부분 직감에 의존하여 운에 모든 것을 맡깁니다.

그렇기에 개인투자자들에게 우리의 아이템이 필요합니다. 주식은 운이다, 운이다 말해도 워렌 버핏 같은 사람이 등장하는 이유는 그들만의 판단 메커니즘이 있기 때문입니다. 데이터를 학습시켜 데이터 속의 보이지 않는 패턴을 학습하는 딥러닝을 이용하여 개인 투자자가 모르는 주식시장의 패턴을 학습한다면 개인의 직감에 의존한 잘못된 투자로 인해 손해를 입는 것을 방지하고 예측된 결과를 토대로 투자의 위험성과 불안감을 줄여 건강한 주식 투자를 할 수 있도록 도와 줄 것입니다.

3) 독창성

기존 시장 아이템

■ 웹페이지

paxnet, stockro, stocker, DC 주식갤러리, 실전주식투자갤러리는 웹 기반 주식 서비스를 제공하는 사이트입니다.

▶ paxnet은 돈을 벌릴 수 있으며 주식의 기본인 관련 정보가 잘 정리되어 있으며 커뮤니티 게시판이 존재합니다. Stockro은 정보 제공의 역할이 paxnet처럼 주식 전반이 아니라 투자 입문의 기본을 알아야 할 것들이 있는 교과서와 같습니다. stocker는 디자인적으로 paxnet보다 좋은 느낌이고 내용은 paxnet의 열화판처럼 유사합니다. Paxnet과 stocker 사이트의 핵심 수익은 대출이며 stockro는 상단부분 광고 화면이 부착되어 있어 페이지뷰, 클릭을 통한 광고수익입니다.

3개 사이트의 공통점은 커뮤니티 기능을 지원하는데 회원제이기 때문에 활성률이 그렇게 높지 않아 보입니다.

▶ 이들과 대조되는 DC 주식갤러리, 실전 주식투자 갤러리인데 DC는 기본적으로 가입하지 않고 글

쓰기가 가능하여 사용자들 스스로 정보를 생성해 공유하며 운영자 측에선 주식 내용을 제공하지 않고 게시판을 관리하는 느낌은 들지 않습니다. 익명성의 보장 때문인지 위 3개의 사이트와 다르게 커뮤니티적 기능이 더 활발하지만 익명성의 부작용(욕설, 정보의 신뢰성) 또한 많은 편입니다. DC의 주요 수익 또한 페이지 곳곳의 광고입니다.

■ 애플리케이션

▶ ‘인공지능추천지식 알리기’ 애플리케이션은 인공지능을 이용하여 추천 종목과 급등, 바닥권 탈출하는 종목 등, 주식 관련 뉴스 또한 제공하며 개인정보를 요구하지 않습니다. 그러나 얼마까지 성장한다고 하는 구체적인 주식 가격 예측은 없으며, 실제 인공지능이라기보다는 거래량 급증, 가격 급증 계산 매크로와 크롤링을 이용해 자동 업데이트 되는 프로그램인 듯합니다. 또한 커뮤니티 기능이 없으며 핵심 수익은 중간중간 광고를 통한 광고 수익입니다.

▶ ‘미래 주가’ 애플리케이션은 오늘 예상 등락, 목표주가 등을 제공하여 예측 서비스를 제공합니다. 또한 채팅방 기능을 통한 커뮤니티적 기능도 제공합니다. 어플 내의 광고가 없으나 개인정보를 요구하며, 유료 회원제를 통해 더 많은 정보, 전문가 의견을 제공하는 것을 핵심 수익으로 하고 있습니다. 단점은 여러 정보를 화면에 꽉 채워서 보여주기 때문에 주식을 아는 사람에게는 상관없을 수 있으나 주식 초보자에게는 답답한 느낌을 줍니다.

시장 아이템을 참고하여 저희가 만들 제품의 특징은

1. 사용자끼리 정보 공유를 위한 플랫폼 제공(익명 커뮤니티)
2. 주식 예측 정보 제공(단순 상승,하락 분류가 아닌 수치예상)

입니다.

4) 필요 기술

- (1) 프로그래밍 언어: Python, HTML, CSS, PHP, MYSQL
- (2) 활용 가능한 오픈소스 라이브러리 설명 및 출처

■ 관련 논문

▶ 텐서플로를 이용한 주가 변동 예측 딥러닝 모델 설계 및 개발

저자 송유정 이종우 숙명여대 it공학과 2017.06

다층퍼셉트론 예측모델을 사용. 일별 시가, 고가, 저가, 종가, 거래량 5개의 지표를 이용하여 5일 후의 종가의 상승 또는 하락을 1또는 0으로 나타내는 프로그램. 약 56% 정도의 성능을 가지며 더욱 성능이 높은 학습모델을 사용하면 발전할 가능성이 존재

▶ 양방향 LSTM 순환신경망 기반 주가예측 모델

저자 주일택 최승호 2018.04

입력값으로 시가,고가,저가,종가,거래량, 은닉층 20개, 손실 함수로 평균제곱근 오차를 사용하여 출력값은 예상 가격으로 양방향 LSTM 구조의 실험 구성, 실험 결과 한방향 LSTM 보다 양방향 LSTM의 성능이 더 우수하다. 정확도는 약 85%

▶ 주식예측 프로그램 : TensorFlow, PyTorch, Keras 중의 하나를 사용해 구현

Tensorflow : 머신러닝을 위한 엔드 투 엔드 오픈소스 플랫폼입니다. 도구, 라이브러리, 커뮤니티 리소스로 구성된 포괄적이고 유연한 생태계를 통해 연구원들은 ML에서 첨단 기술을 구현할 수 있고 개발자들은 ML이 접목된 애플리케이션을 손쉽게 빌드 및 배포할 수 있습니다.

<https://www.tensorflow.org/>

PyTorch : Python을 위한 오픈소스 머신 러닝 라이브러리로서 Facebook 인공지능 연구팀이 개발하였습니다. Torch를 기반으로 하며, GPU사용이 가능하기 때문에 빠른 속도로 연산이 가능합니다. Tensorflow에 비해 코드가 간결하고 낮이도가 낮으며, 직관적인 특징이 있다.

<https://pytorch.org/>

Keras : 파이썬으로 작성된 오픈소스 신경망 라이브러리입니다. MXNet, Deeplearning4j, 텐서플로, Microsoft Cognitive Toolkit 또는 Theano 위에서 수행할 수 있습니다. 딥 신경망과의 빠른 실험을 가능케 하도록 설계되었으며 최소한의 모듈 방식의 확장 가능성에 초점을 둡니다. ONEIROS(Open-ended Neuro-Electronic Intelligent Robot Operating System) 프로젝트의 연구적 노력의 일환으로 개발되었으며 주 개발자이자 유지보수자는 구글의 엔지니어 Francois Chollet입니다

<https://keras.io/>.

▶ 웹 사이트 : 사이트 제작 및 DB 관련된 내용은 생활코딩 참조

생활코딩은 일반인에게 컴퓨터 프로그래밍을 알려주는 것을 목적으로 하는 비영리 교육 프로젝트입니다. 교육은 이고잉이 만든 오픈튜토리얼스에서 제공됩니다. 웹서비스를 중심으로 설명하는 '생활코딩', 중장년을 대상으로 IT 활용법을 다루는 '효도코딩' 등으로 구성됩니다. 프로그래머 이고잉(egoing)이 시작했고 2011년부터 오프라인 강의를 병행했습니다.

<https://opentutorials.org/course/1>

5) 월별 계획 설정

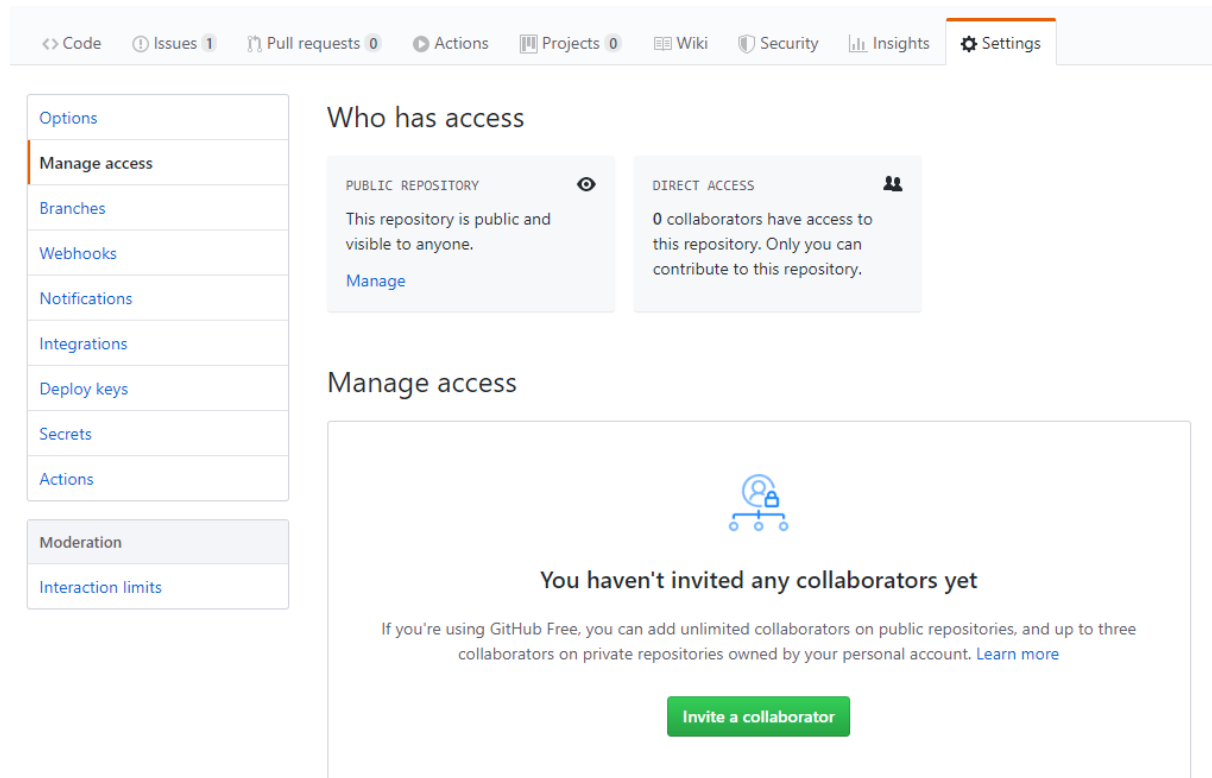
월	계획
4월	MLP 모델 구현, 크롤링 기능 구현, 웹과 DB 설계 및 웹과 DB 연결, LSTM 학습
5월	MLP 모델 테스트 및 디버깅, 단방향 LSTM모델 제작
6월	웹에 커뮤니티 기능 추가, 모바일 앱 디자인 및 초기 버전 구현, 단방향 LSTM 테스트
7월	양방향 LSTM 학습, 부족한 부분이나 수정사항 보충
8월	기존 모델 양방향 LSTM으로 교체, 모바일 앱 완성, 뉴스 분석 사전 생성
9월	부족한 부분이나 수정사항 보충, 앱(구글 스토어) 및 웹 릴리즈
10월	앤드류저 피드백, 트래픽 검사 등을 통한 수정사항 도출 및 수정

6) 주차별 목표 설정

주차	주차별 계획 및 담당자
04.13-04.19	<ul style="list-style-type: none"> - 프로젝트 설계(웹 디자인, 필요한 클래스, 인터페이스)(김경중) - 크롤링 기능 구현, 코드 리뷰(전부) - DB 기능 구현, 코드 리뷰(전부) - 크롤링 기능 수정(김진용) - DB 기능 수정(김현재) - 깃허브 사용법 학습(김진용) - 크롤링과 DB 기능 통합 테스트 후 결과 공유(김현재)
04.20-04.26	<ul style="list-style-type: none"> - 테스트 결과 리뷰 및 수정사항 도출 후 수정(김현재) - MLP 모델 구현(김경중) - 모델 인터페이스 및 DB 인터페이스 구현(김진용) - MLP 모델 및 LSTM 모델 학습(김현재, 김진용)
04.27-05.03	- 중간 시험기간 대비
05.04-05.10	<ul style="list-style-type: none"> - MLP 모델 학습 및 결과 저장(김진용) - 크롤링->DB->Python(MLP)->DB->Web 프로세스 통합 테스트(김현재) - 전체 결과 리뷰 및 수정사항 도출(김경중)
05.11-05.17	<ul style="list-style-type: none"> - 모바일 앱 디자인, 웹 게시판 레이아웃 설계(김경중) - 앱 개발 학습(김진용) - 웹 게시판 기능 학습(김현재)
05.18-05.24	<ul style="list-style-type: none"> - 앱 개발(김진용) - 웹 게시판 기능 개발(김현재) - LSTM 구현(김경중)
05.25-05.31	<ul style="list-style-type: none"> - 앱 개발(김진용) - 웹 게시판 기능 개발(김현재) - LSTM 구현(김경중)
06.01-06.07	- 기말 시험기간 대비
06.08-06.14	<ul style="list-style-type: none"> - 구현 된 모든 기능 통합 테스트 및 수정사항 도출(전부) - 발표용 PPT 및 자료 정리(전부)
06.19, 06.26	프로젝트 중간결과 발표

<참고자료>

1. Sourcetree를 사용하는 Github 튜토리얼 : <https://milooy.wordpress.com/2017/06/21/working-together-with-github-tutorial/>
2. 저장소에서 collaborator 추가 메뉴 : Settings > Manage access > Invite a collaborator



3. Git/Github의 개념 및 Command line 환경에서의 Git/Github 사용 : e-campus의 3주차 파일 참고