



KT&G, LG생활건강 주가 예측 프로젝트

one page proposal

이채림

학습 모델 RNN / LSTM

시계열 데이터인 주가를 예측하기 위해

딥 러닝 모델 중 순환 신경망 RNN(Recurrent Neural Network) 과

RNN의 한 종류인 LSTM (Long Short-Term Memory)을 사용

학습 기간

LG생활건강 (2002-01-01~Now)

KT&G (2000-01-01~Now)


진행 과정

- ☐ Finance Data Reader 사용
- ☐ 수업 내용 복습 및 심화 학습
- ☐ KT&G, LG생활건강 재무데이터 크롤링
- ☐ RNN, LSTM 딥러닝 머신으로 학습하여 주가 예측
- ☐ 예측 날짜와 변수를 조정해가면서 가장 잘 맞는 모델을 추출
- ☐ 프로젝트 결과 분석 및 비교

활용

Welcome to Python.org


The official home of the Python Programming Language

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



pandas

pandas is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language.

 <https://pandas.pydata.org/>



Project Jupyter


JupyterLab is the latest web-based interactive development environment for notebooks, code, and data. Its flexible interface allows users to configure and arrange workflows in data science, scientific computing, computational journalism, and machine learning. A

 <https://jupyter.org/>



Matplotlib

Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. Matplotlib makes easy things easy and hard things possible. Try Matplotlib (on Binder) Be sure to check the Users guide and the API docs. The full text

 <https://matplotlib.org/>



NumPy

Why NumPy? Powerful n-dimensional arrays. Numerical computing tools. Interoperable. Performant. Open source.

 <https://numpy.org/>

scikit-learn

"We use scikit-learn to support leading-edge basic research [...]" "I think it's the most well-designed ML package I've seen so far." "scikit-learn's ease-of-use, performance and overall variety of algorithms implemented has proved invaluable [...]."

 <https://scikit-learn.org/>



GitHub - tensorflow/tensorflow: An Open Source Machine Learning Framework for Everyone

Documentation TensorFlow is an end-to-end open source platform for machine learning. It has a comprehensive, flexible ecosystem of tools, libraries, and community resources that lets researchers push the state-of-the-art in ML and developers easily build and deploy ML-powered applications.

 <http://github.com/tensorflow/tensorflow>

tensorflow/ tensorflow

An Open Source Machine Learning Framework for Everyone

 3k Contributors 184k Used by 164k Stars 87k Forks

Keras: the Python deep learning API

Keras is an API designed for human beings, not machines. Keras follows best practices for reducing cognitive load: it offers consistent & simple APIs, it minimizes the number of user actions required for common use cases, and it provides clear & actionable error messages. It also has extensive documentation and developer guides.

 <https://keras.io/>

FinanceDataReader 사용자 안내서

금융 데이터를 다루는데 가장 기본이 되는 데이터는 거래소별 전체 종목 코드와 가격 데이터 이다. pandas-datareader 는 잘 구성된 시계열 데이터 수집 라이브러리로 사용이 간편하고 다양한 시계열 데이터를 수집할 수 있다는 장점이 있다. (현재 버전 : pandas_datareader 0.6.0) 하지만, 거래소별(KRX, NASDAQ, NYSE 등) 전체 종목 코드(ticker symbol)를 가져오는 기능이 없으며, 야후 파이낸스가 더 이상지원되지 않고(deprecated), 구글 파이낸스는 <https://financedata.github.io/posts/finance-data-reader-users-guide.html>

<https://github.com/FinanceData/FinanceDataReader>

전자공시시스템

 <https://dart.fss.or.kr/>