

[PA01] Linear & Logistic Regression

올해 부산대 정컴 4 학년이 되는 산지니는 머신러닝 수업을 들으며 주가 예측 아이디어를 얻었고 벼락부자가 될 생각을 한다. 산지니는 과거 주가 데이터를 활용해 주가 예측 모델을 개발하고 튜닝하려고 한다. 요즘 핫하다는 머신러닝을 활용해 주가 예측을 하여 산지니를 도와보자!

“주식 시장은 인내심이 없는 자로부터 인내심이 많은 자에게로 돈이 넘어가도록 설계되어있다. - 워렌 버핏”

Data: (2023.07.03 ~ 2023.12.22 의 주식 데이터) (*cp949 로 csv 파일 인코딩하여 로드)

- stock_list.csv (종목 리스트) – 364 rows x 3 cols
- stock_values_linear.csv (linear regression 을 위한 주식 가격 정보) – 364 rows x 125 cols
- stock_values_logistic.csv (logistic regression 을 위한 주식 가격 정보) – 364 rows x 125 cols

X : (월 ~ 금) * 24 주간

y : (다음주 월 ~ 금) * 24 주간

y_0 : 다음주 월요일, y_1 : 다음주 화요일, y_2 : 다음주 수요일, y_3 : 다음주 목요일, y_4 : 다음주 금요일

- 이번주 월~금요일의 패턴을 학습해 다음주 월요일~금요일을 각각 예측하는 모델을 생성
- 이 과정을 모든 종목(364 개)에 적용

Q1. Stock list 에 있는 주식 3 개의 다음 주의 가격을 Linear regression 을 통해 예측하라.

Q2. Stock list 에 있는 주식 3 개의 다음 주의 유형을 Logistic regression 을 통해 예측하라.

[0: 유지, 1: 올라가는 추세, 2: 내려가는 추세 (전날 가격 기준)]

제출 형식

1. 코드 구현 및 실행 결과가 저장된 ipynb 파일
2. 그림, 표 등을 포함하는 보고서를 작성하여 제출하나, 분량은 그림, 표 등을 제외한 텍스트 기준 최소 워드 2 장 이상의 분량으로 제출
3. 각각의 알고리즘 성능 평가 및 결과 분석
4. (Optional) 현재 머신러닝 수업을 수강중인 동료에게 도움을 받았을 경우, 동료의 학번/이름 (추후 표절 시비를 방지하기 위함)