

Unit 2

기계학습응용 :
주식가격 예측

4주차. 기계학습

학습 내용

- 기계학습이란?
 - 데이터 마이닝, 기계학습, 인공지능, 딥러닝
- 주식가격예측

학습 목표

- 기계학습의 개념과 방법들을 이해하고 설명할 수 있다.
- 기계학습이 주식가격 예측에 어떻게 활용되고 있는지 알 수 있다.

주식가격 예측

☑ 기본적 분석 vs 기술적 분석

재무분야에서의 주식가격 예측 방법 2가지

기본적 분석
(Fundamental Analysis)

기술적 분석
(Technical Analysis)

주식가격 예측

☑ 기본적 분석 vs 기술적 분석

	기본적 분석	기술적 분석
분석 데이터	재무제표, 경제 통계	과거 데이터
분석기법	내재가치 도출	시계열분석 (Serial dependence)

- 데이터 간의 serial dependance, 시계열의 의존성을 파악해서 모형화

주식가격 예측

☑ 기본적 분석 vs 기술적 분석

	기본적 분석	기술적 분석
시장의 효율성	Semistrong-form Efficiency	Weak-form Efficiency

➤ 과거 및 현재 데이터 모두 활용

➤ 데이터셋이 상대적으로 적음

주식가격 예측

☑ 기본적 분석 vs 기술적 분석

	기본적 분석	기술적 분석
투자 전략	내재가치와 시장가치 차이 활용	주가와 추세와의 차이 활용
	시장가치 대비 저평가/고평가 판단	모형에 의해 추정된 추세와 실제 주가 사이의 괴리

» 추세 = 균형가격

» 현재 시장이 곧 균형으로 돌아올 것이라고 가정

» 괴리를 투자전략에 활용

주식가격 예측

☑ 기본적 분석 vs 기술적 분석

	기본적 분석	기술적 분석
투자 성향	가치 투자	예측, 심리투자
투자 기간	중장기	단기

➤ 단기적인 data frequency가
아니기 때문에

주식가격 예측

☑ 기본적 분석 vs 기술적 분석

Daily data를 많이 활용


	기본적 분석	기술적 분석
한계점	재무제표 데이터 4회/년	단기투자자에 집중

➤ 통상 실증분석에서는 연 데이터 사용

한계점에 있지만 재무분야에서 전통적으로 많이 활용되는 방법

주식가격 예측

☑ 주가예측

 여러분에게 과학기술,
새로운 정보가 있다면,
그 정보나 기술이 돈과 관련 있다면 ?

 최신 분석기법, 투자기법을 응용해
초과수익을 얻으려 노력

주식가격 예측

☑ 주가예측

주식시장은 효율적

- » 모든 정보가 주식가격에 신속하게 반영이 되기 때문에 초과수익률을 얻기 힘들
- » 시장이 모르는 정보나 기술습득은 어려움

한계점에 있지만 재무분야에서 전통적으로 많이 활용되는 방법

주식가격 예측

☑ 주가예측

주가 pricing 모형

- » 특별한 이론 없이 요인 모형이라는 factor model에 근거한 pricing 모델이 응용
- » 완성된 pricing theory가 아닌
아직까지 **진행중인 pricing theory**



주가 예측 또는 주가수익률 예측은 매우 힘든 과제이다!

금융AI를 배우는 이유는 **과학적 투자**를 위한 것

주식가격 예측

☑ 주가모형에 대한 이해

성공적인 주가예측 모형이란?

최소한의 입력데이터와 **복잡하지 않은** 모형

- » 이론이 복잡하고 많은 데이터를 활용할 경우
투자 측면에서는 부정적
- » 핵심변수, 최소한의 설명변수로 최대한
종속변수를 설명하는 것이 좋은 모형



복잡성이 꼭 모형의 성과에 긍정적인 것은 아님을 명심하라.

주식가격 예측

☑ 주가 pricing의 발전

재무이론의 주가 pricing 발전

시계열

횡단면

- » 주가의 시계열 데이터는 변동이 심함
- » 수가 수준 데이터 : stationarity 하지 않음
- » 수익률 데이터 : stationarity에 가까움



stock return에 대해서 재무이론이 발전

주식가격 예측

☑ 주가 pricing의 발전



Eugene Fama

결합가설문제
(Joint hypothesis problem)

“pricing model은 완벽하지 않기 때문에
시장 효율성을 실제로 테스트 하기는 힘들다.”

➔ 한계가 있지만 event study등을 통해서
마켓 효율성을 계속 실증 분석 중

주식가격 예측

☑ 주가 pricing의 발전

벤치마크 theory

90년대 초반
3요인 모형



Fama and French, 2015년
5요인 모형

➤ 벤치마크 또는 출발점처럼
받아들여지고 있음

아직까지도 stock return에 관한 pricing theory는 진행형

주식가격 예측

☑ 주가 pricing의 발전



(+) 유동성 요인, 모멘텀 요인 및 새로운 요인을 계속 탐색중

주식가격 예측

☑ AI분야의 주가예측

기계학습, AI, 딥러닝의 도입

데이터와 알고리즘에 근거한
pricing theory 시도

- » 기존의 재무 분야하고 완전히 다른 분야에서 주가 예측 시도
- » 데이터사이언스 측면의 기계학습 접근은 아주 새로운 방법
- » 기존 이론과 관련성도 많지만 새로운 방법론을 제시하고 있음

주식가격 예측

☑ AI분야의 주가예측

Data-driven

non-parametric approach

주식가격 예측

☑ AI분야의 주가예측

기계학습, AI, 딥러닝의 도입의 단점

- 1 많은 controlling parameters가 필요
- 2 stable solution을 찾기 어려움
- 3 과적합(over-fitting) 문제 자주 발생



feature engineering

- 핵심적인 feature를 찾는게 필요

많은 진보를 보이고 상당히 빠른 속도로 발전하고 있는 분야

주식가격 예측

☑ AI분야의 주가예측

Option Pricing Model(OPM)

(Black, Scholes, and Merton)

주식 pricing에는
활용이 안됨

- » Stock price 대신 **stock return**으로 pricing 시도
- » 기존 모형으로는 일정 부분 설명이 안되는 부분 발생
= **Anomaly**
- » Anomaly를 줄이기 위해 **새로운 factor** 탐색



Stock return model 발전하는 과정

주식가격 예측

☑ AI분야의 주가예측

머신러닝, AI분야의 기초 모형

- » ARIMA, AR 모델 + Ensemble method
- » Support vector machine (SVM)

벤치마크 대비
성과 좋음

주식가격 예측

☑ AI분야의 주가예측

딥러닝 모형

- » **LSTM RNN** (long short-term memory recurrent neural network)
- » 텍스트마이닝 기술 발전
 - **Sentiment Analysis** : 감성분석

주식가격 예측

☑ Sentiment Analysis

Sentiment Analysis

뉴스가 텍스트마이닝 기술을 통해서 데이터로 추가되는 기술



텍스트마이닝

- weighting mechanism
- 모두 주가에 관련 있지 않음

핵심정보
selection
extraction

다층 레이어 도입 등이 주가 예측력의 향상에 도움

주식가격 예측

☑ Sentiment Analysis

핵심 feature 추출 방법

Multi-filters neural network (MFNN)

Hybrid attention network (HAN)

➤ attention mechanism + natural language processing

주식가격 예측

☑ 독립성분분석

독립성분분석

(ICA: independent component analysis)

효과적으로 feature를 추출하는 방법

- » 기존에 재무이론에서 시도되지 않았던 방법
- » factor 이론과 관련이 있지만 더 advanced된 방법론



주가 예측 활용에 뛰어난 성과



이론적으로 복잡하다는 단점

주식가격 예측

☑ 독립성분분석

독립성분(independent component)

관찰 데이터의 hidden information

선형 ICA

» 관찰 데이터가 독립성분의 **선형결합**으로 생성

주식가격 예측

☑ 독립성분분석

독립성분(independent component)

비선형 ICA

- » 관찰 데이터가 독립성분의 **비선형결합**으로 생성
 - 독립성분의 Source나 mixing procedure에 관한 사전지식을 요구하지 않음

주식가격 예측

☑ 독립성분분석



➤ 나머지는 노이즈

➤ 선형/비선형

➤ 주가수익률 복원

➤ IC들로 결합된 주가수익률 = **균형**

➤ **균형수익률**과 **시장수익률** 사이의 괴리 활용

➤ 다른 접근법에 비해 **성과가 좋음**

주식가격 예측

☑ 가격이론과 시장 효율성

? advanced pricing theory나 알고리즘이 시장효율성에 어떤 영향을 미칠까



- » 엄청난 노력, 정보분석
- » 알고리즘을 발견, 활용
- » 단기적으로 추가수익률 가능



시장에 반영되지 않은 가격이론, 정보나 기술일 경우



주식가격 예측

☑ 가격이론과 시장 효율성

No arbitrage 가정

추가수익 찬스

경쟁심화

Arbitrage profit $\rightarrow 0$

No arbitrage = equilibrium

“ 시장효율성 달성 ”