

초단타매매와 데이터를 이용한 퀀트 트레이딩

YONGJIN KWON

QUANTITATIVE ANALYST

소개

권용진 (Yongjin Kwon)

Carnegie Mellon Univ. Computer Science / Applied Mathematics

前 시카고 러시 대학 병원 환자 데이터베이스 개발자

前 피츠버그 로봇공학 연구소 뇌파 인공지능 연구원

前 뉴욕 뱅크오브아메리카 메릴린치 초단타 옵션 퀀트 애널리스트

前 뉴욕 타워리서치 캐피탈 외환 초단타 트레이딩 퀀트 애널리스트

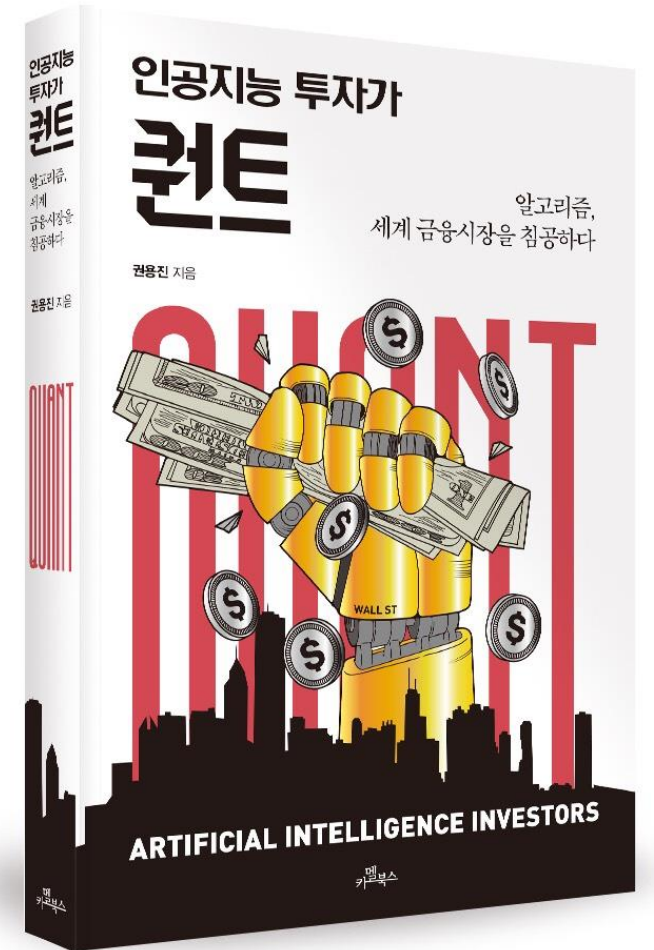
現 엔트로피 트레이딩 그룹 대표 / 포트폴리오 매니저

소개

인공지능 투자가 퀀트
(리디북스, Yes24, 교보 영풍 등 오프라인 서점)

www.quant-story.com

ykwon@entropy-trading.com



소개 – Tower Research Capital



세계 최대의 초단타 퀀트 트레이딩 헤지펀드

90%의 트레이더가 이공계 출신
(물리학, 전자공학, 수학, 컴퓨터공학 등)

다수의 슈퍼컴퓨터 보유

병렬처리, 머신러닝, 신호처리 등 고급 엔지니어

뉴욕 / 시카고 / 런던 / 인도 / 싱가포르 / 홍콩 오피스

오늘의 강연

퀀트의 정의와 현재 트렌드

초단타매매와 시장 미시구조, 실제 알고리즘 전략들

데이터를 이용한 투자 전략 수립 및 시뮬레이션 소개

Q & A

퀀트?

고도의 수학, 통계지식을 이용해서 투자법칙을 찾아내고 컴퓨터로 적합한 프로그램을 구축해서, 이를 토대로 투자를 행하는 사람을 일컫는 말이다.

- 한국 경제 연구소

그러나 다변화된 퀀트

퀀트 트레이더

퀀트 애널리스트

퀀트 인베스터

모델 검증 퀀트

퀀트 개발자

리스크 관리 퀀트

알고리즘 트레이더

Rule of Thumb

1.

퀀트와 알고리즘 -> 미래와 가격 예측 (X)

확률 상으로 유리한 상황을 연구 ->
반복된 같은 상황으로 큰 수의 법칙에 의한 우위를 실현(O)

2.

돈을 자동으로 벌어드주는 프로그램 제작 (X)

의사 결정 과정을 수치 / 데이터화해서 자동화 -> 심리 편향 없이 실행(O)

우위?

통계 분석을 통해서 확률상으로 높은 사건을 찾는 것

예시)

스포츠 분석

일기 예보

포커

Trend

- 전산화 – Arbitrage
- 모델화 – Statistical Arbitrage, Factor Analysis, Derivative Pricing
- 고속화 – High Frequency Trading, Electronic Market Making
- 데이터화 – Pattern Recognition, Hidden Markov Model
- 인공지능화 – Natural Language Processing, Machine Learning

전산화

-거래소 및 정보가 전산화 되면서 널리 이용된 전략들

Arbitrage (무위험 차익거래)

같은 물건 혹은 구성이 비슷한 물건들의 가격이 다를 때 차액을 버는 방식

-유로화가 미국에서 1.1달러, 캐나다에서 1.2달러 일 경우

-일년 이자가 5%인데 100만원 입금시 일년뒤 110만원 주는 상품이 있는 경우

Index Arbitrage

-인덱스(여러가지 상품의 그룹)의 가격과 각각의 상품 가격이 다를 경우

모델화

어떤 현상에 대한 공식이나 확률 분포를 만들고 이를 이용해 수익

-Derivative Pricing

보험과 같은 확률 프로세스가 포함된 상품의 가격을 계산해서 설계해주고, 프리미엄을 붙여서 판매

이세돌-알파고 이벤트 10억 보장 상품 가격

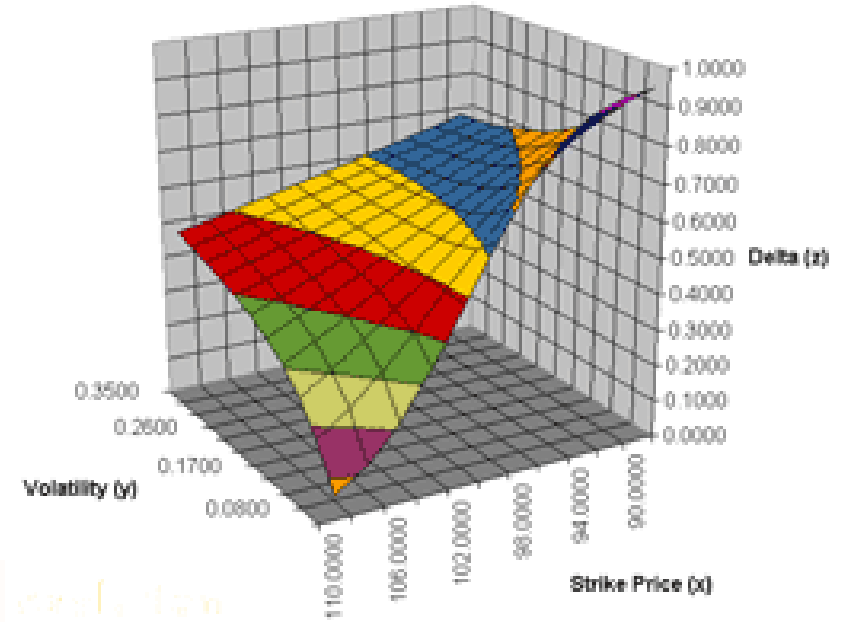
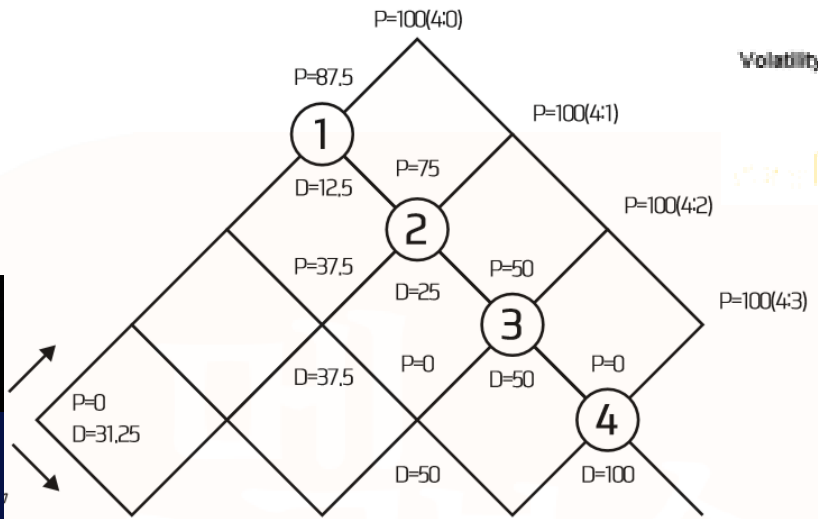
수출 업자의 환율 보장 상품 가격

날씨 보험 상품 가격

Derivative Pricing

알파고 vs 이세돌 문제

- 7전 4선승제
- 단일 경기 베팅 가능
- 승률 50:50
- 이세돌 승리시 10억 지출

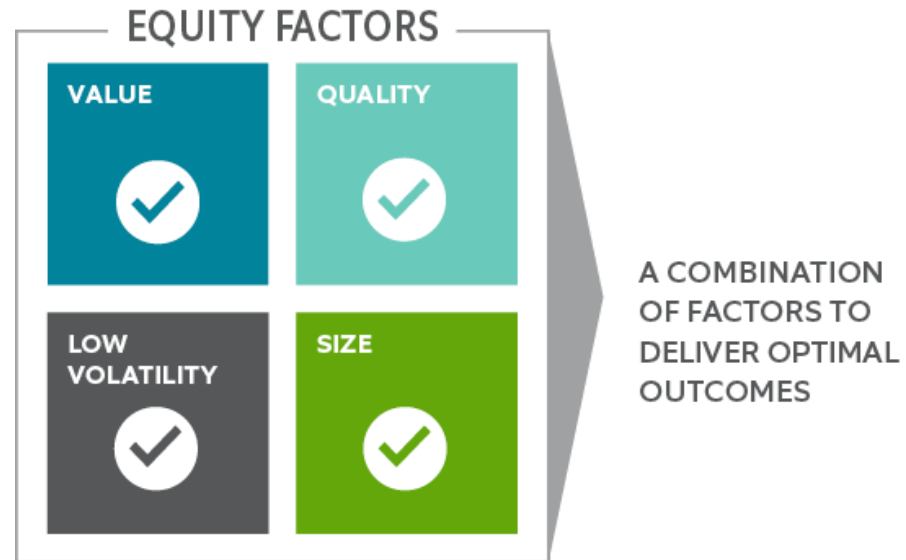


모델화

-Factor Analysis

자산 가격의 움직임에 영향을 주는 요인을 분석

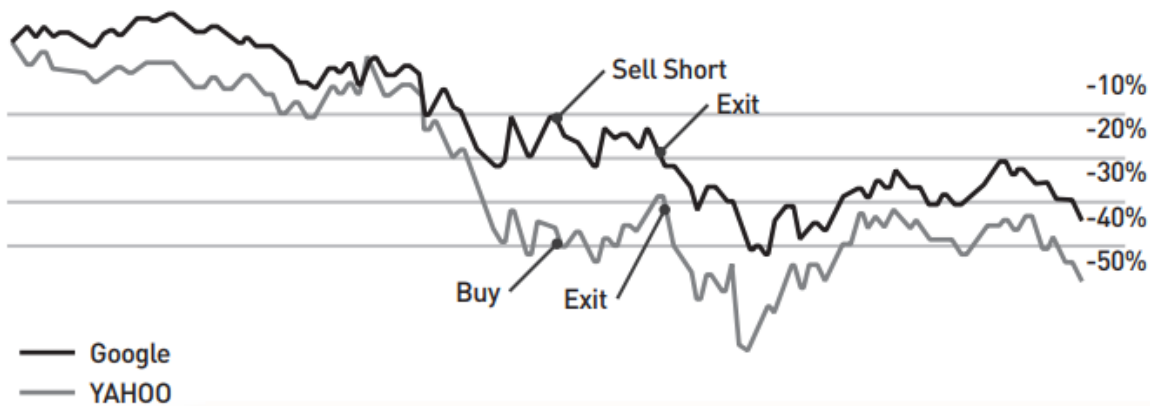
이를 요인 별로 분해한 뒤에 움직임 예측



데이터화

-Statistical Arbitrage

통계적으로 움직임이 관계가 있는 상품들의 관계에 베팅
두 상품으로 하는 Pair Trading, 그룹으로 하는 Basket Trading 등
퀀트 트레이딩에서 가장 유명하고 인기있는 전략



데이터화

Pattern Detection

노이즈 필터링

Hidden Markov Model

음성 인식 기술

PCA 분석

몬테카를로 시뮬레이션

다양하고 새로운 방식을 동원해서 주기적으로 일어나는 패턴을 찾아냄

데이터화

-히든 마코브 모델 (HMM)

숨겨진 패턴을 보이는 데이터를 이용해서 찾는 기술
음성 인식 등에서 많이 사용함

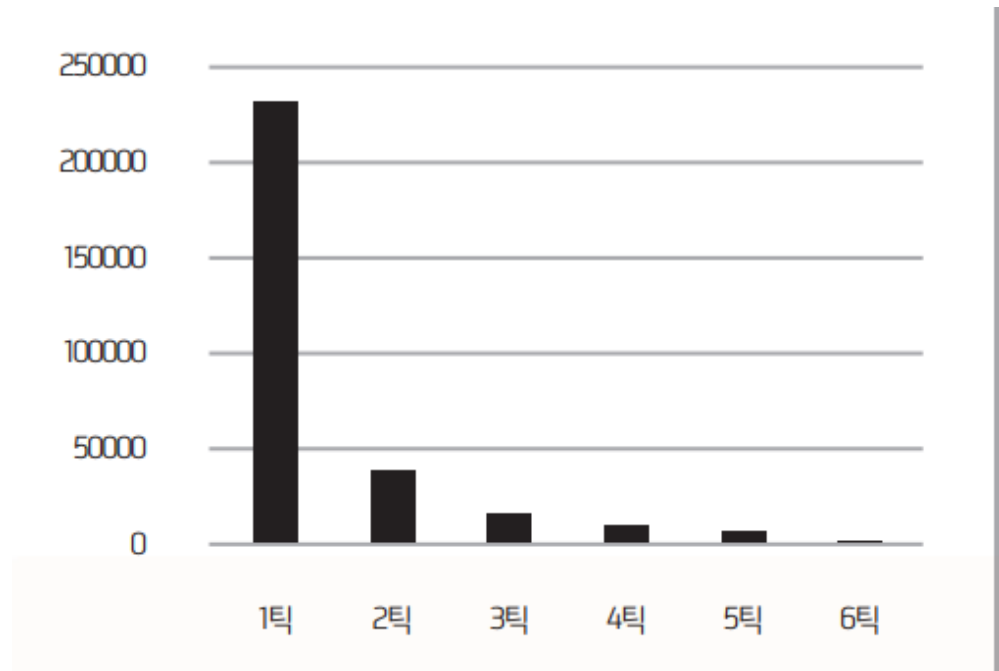
고속화

-HFT / Market Microstructure Modeling

호가 분석해서 움직임 예측하는 전략

주문을 어떤식으로 분포시키는가 등의 중요

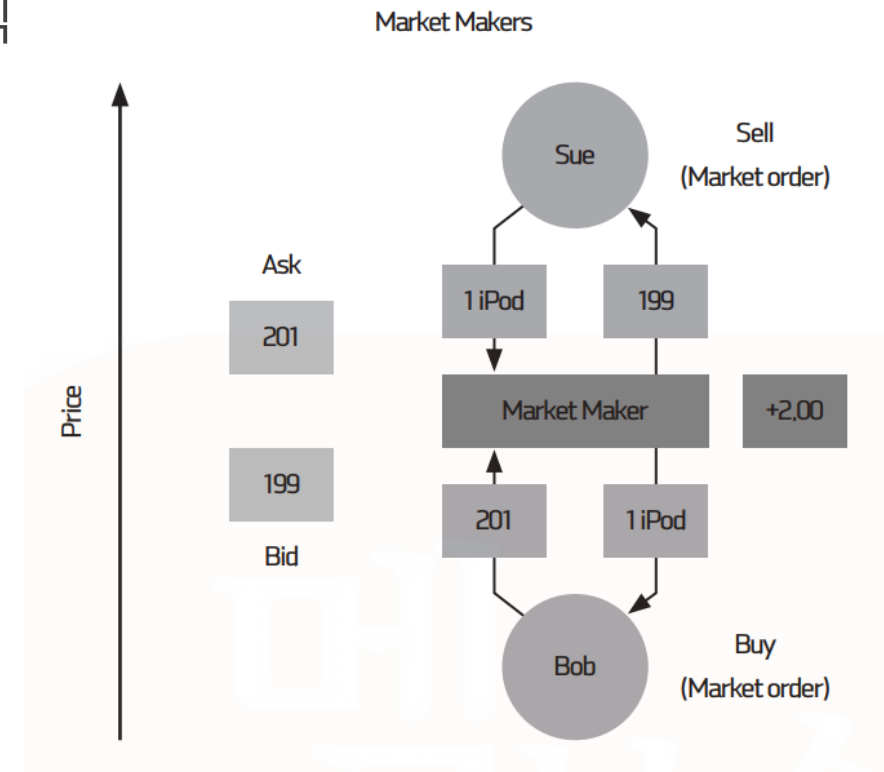
건수	수량	가격	수량	건수
5	32	102		
18	89	101		
112	1042	100		
		99	13	5
		98	45	22
		97	829	65
		96	157	38
Offer			Bid	



고속화

-Electronic Market Making

사는 주문과 파는 주문을 동시에 올려 놓고 차액을 얻는 방식
팔려는 사람과 사려는 사람의 유동성 공급으로 수익
재고가 쌓이면 손실이 생기므로 재고 관리 알고리즘이 핵심
빠른 속도와 네트워크 지연 시간 최적화가 필요



고속화

-Algorithmic Trading

자신의 거대 주문을 시장 충격을 최소화 시키면서 시장에 거래를 하는 방법

VWAP, TWAP, Implementation Shortfall

시티은행 - 슬라이서

크레딧 스위스 - 게릴라, 스나이퍼

고속화

-속도 전쟁

물리적 거리 최적화

전기 및 회로 전문가 (FPGA)

병렬 처리 엔지니어 (그래픽카드, 슈퍼컴퓨터)

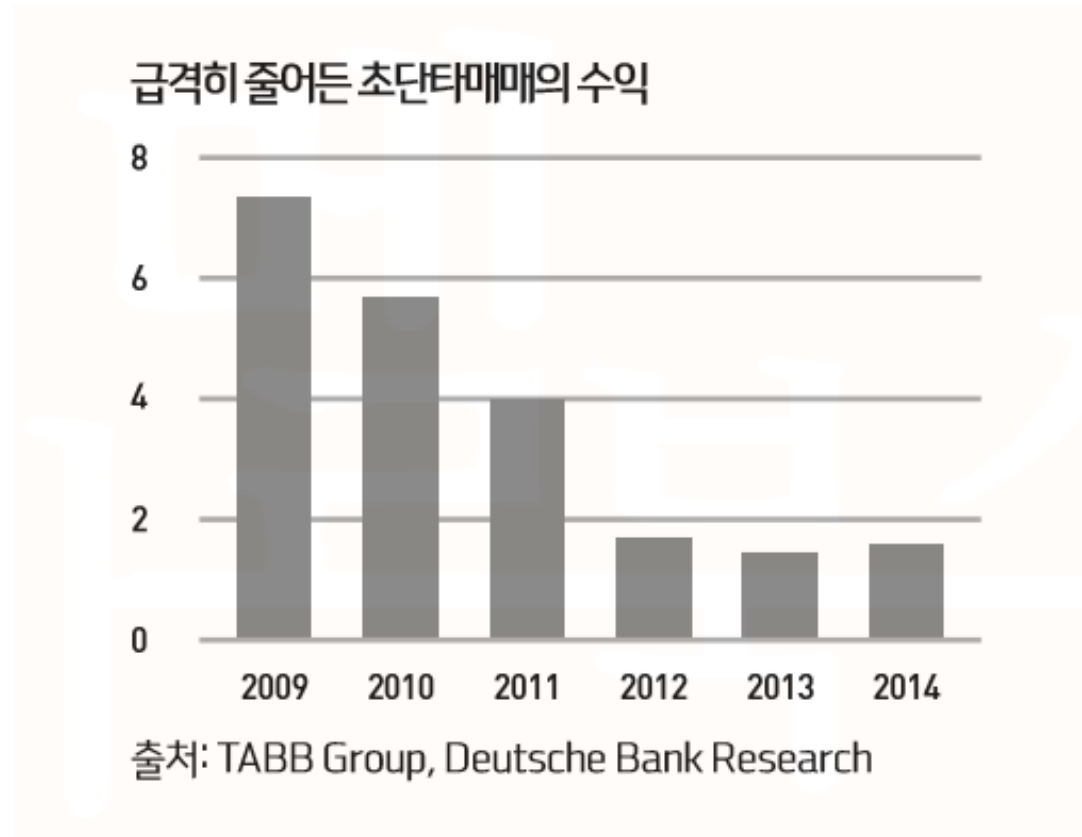
분산 및 고속 데이터베이스 처리

원자 시계 / 마이크로파 / 레이저



고속화

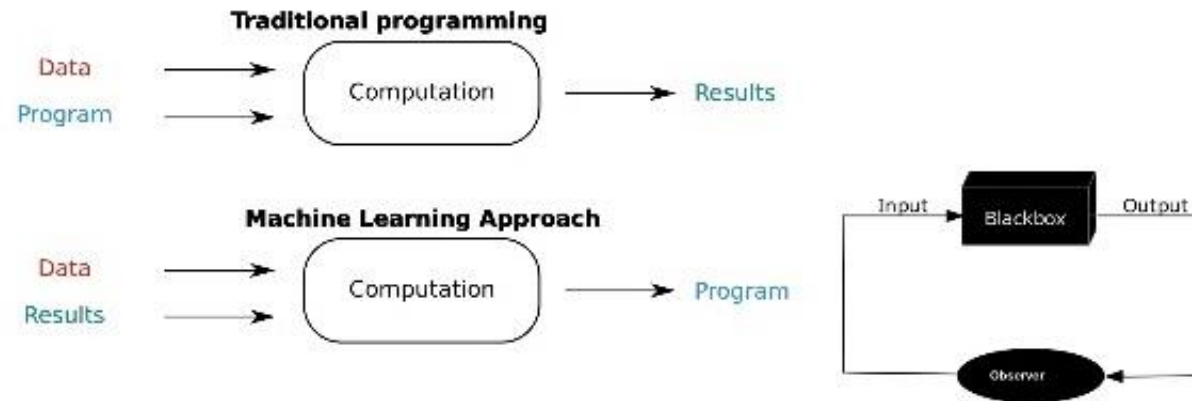
-초단타매매의 한계 봉착



인공지능화

기존에는 모델링 능력이 중요
인공지능(머신러닝)의 등장으로 모델링 능력의 중요도 하락

- 블랙박스 접근 vs 룰기반



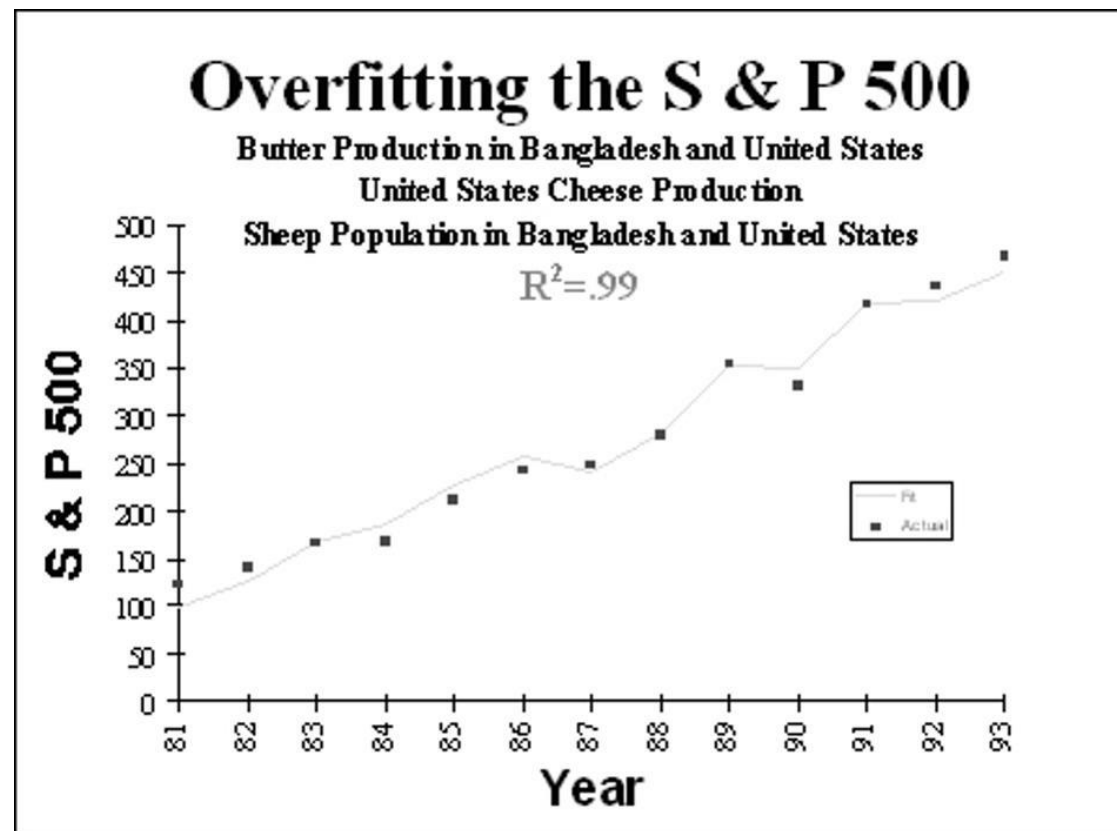
인공지능화

인공 신경망, Decision Tree 등
Classification을 이용한 주가 예측 -> 실패

이유 :

지나친 Overfitting
Feature의 복잡함
Economical Reasoning 부족

인공지능화



인공지능화

주가 예측 보다는 머신러닝 기술을

기존의 데이터 중에 분석이 어려웠던 패턴에 이용

- NLP를 이용한 뉴스 / 소셜 분석
- 음성 처리를 이용한 음성 데이터 분석
- 각종 비정형 데이터 / 알터네이티브 데이터 이용

인공지능화

뉴스 NLP



형태소분석(morphological analysis)

도근 분리, 어간 추출, 품사 부착, 색인, 벡터화

구문분석(syntax analysis, parsing)

문장 경계 인식, 구문분석, 공기어, 개체명 사전 구축(P,OT, 수치, 외국어 한글 표기), 개체명 인식

의미분석(semantic analysis)

대응어 해소(대명사, 두문자어, 약어, 수치), 의미 중의성 해결(동명어인, 이명동언)

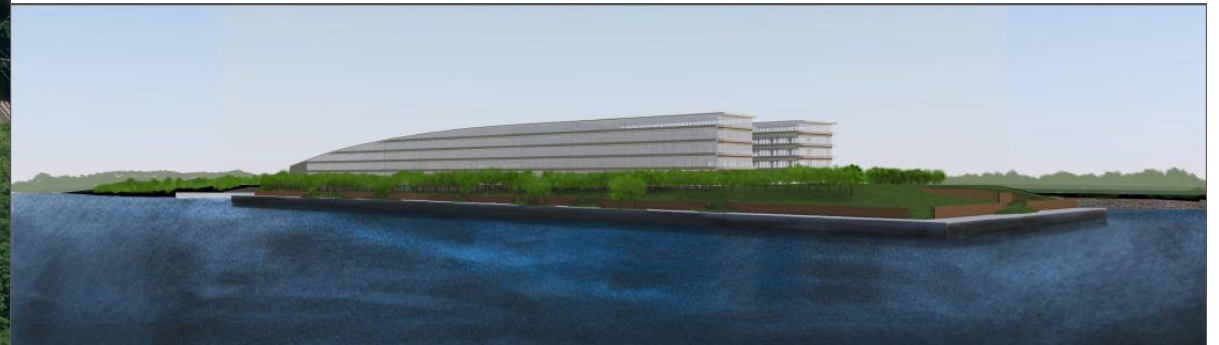
담론분석(discourse analysis)

분류, 군집, 중복, 요약, 가중치, 순위화, 평판분석, 감성분석, 토픽 모델링, 이슈 트래킹, 복합논증분석

인공지능화



Renaissance Technologies – 인공위성 사진 분석의 선구자



Bridgewater Associate – 글로벌 매크로 다양한 인공지능 시도

인공지능화



인공지능화

발전하는 머신러닝 알고리즘에 따른 트레이딩 알고리즘 변화

- 강화 학습 (Reinforcement Learning)
- 트레이더의 행동과 판단기제 자체를 학습
- 카오틱 시스템인 주식 시장 자체를 모델링
- 빅 데이터를 이용한 거대 매크로 경제 모델 정립

최신 트렌드

한계에 다다르는 속도 경쟁 -

슈퍼 컴퓨터, 마이크로파, 레이저, FPGA, 병렬처리

기존 초단타 패턴들의 레드오션화

알터네이티브 데이터 / 데이터 과학의 도입

수학 및 물리학자 -> 데이터 과학자

NLP / 비정형 데이터 전문가

리서치 시스템 개발의 중요성 부각, 퀀트 개발자의 중요도 상승

초단타 매매 / 알고리즘 트레이딩

정의

정해진 규칙을 가진 프로그램이 1초에 적게는 몇 번, 많게는 수백 번의 거래를 하는 전략. 시장 미시 구조 분석과 밀접한 관계를 가지며 작지만 확실한 수익을 얻는 것을 목표로 한다.

시장 미시 구조 vs 재무 분석

- 시장 미시 구조 (Market Microstructure)
 - 주가가 형성되는 과정을 연구
 - 지정가 주문 / 시장가 주문 / 취소 주문 / 거래량 / 스프레드
 - 주문의 흐름 (Order flow)를 통한 수익 실현
 - 틱 데이터 (주문 데이터)를 이용

- 재무 구조 분석
 - 주가를 기준으로 기업을 평가
 - 포트폴리오 이론, 펀더멘탈, 일물일가 등
 - 저평가 된 기업 발굴, 혹은 장기적 시장 흐름을 분석
 - 과거 주가 데이터를 이용

시장 미시 구조 특징

- 효율적 시장이 아님
- 정보의 불균형 존재
- 단기적 패턴
- 알고리즘 트레이딩은 불균형을 해소시킴과 동시에 수익을 실현

호가창

009150 20% 삼성전기			
76,600 ▼		200	-0.26%
316,465		77.66%	
증감	76,600	76,500	24,191 0.42%
매도잔량	11,820	77,500	KOSPI200
	8,308	77,400	76,300 시
	10,371	77,300	77,100 고
	9,350	77,200	76,100 저
	12,752	77,100	76,800 기준
	8,702	77,000	99,800 상
	6,178	76,900	53,800 하
	4,007	76,800	254 비용
	6,419	76,700	76,600 예상
	1,713	76,600	96 수량
			현재가 76,600 ▼200 -0.26%
76,600	10 ▲	76,500	60
76,600	87	76,400	230
76,600	100	76,300	1,705
76,600	5	76,200	9,617
76,600	1	76,100	15,816
76,600	1	76,000	11,636
76,600	1	75,900	11,114
76,600	16,196	75,800	15,314
76,400	9	75,700	13,025
76,400	24 ▼	75,600	10,835
			매수잔량
	79,620	16:00:01	89,412
	5	시간외	

시장 참여자 종류

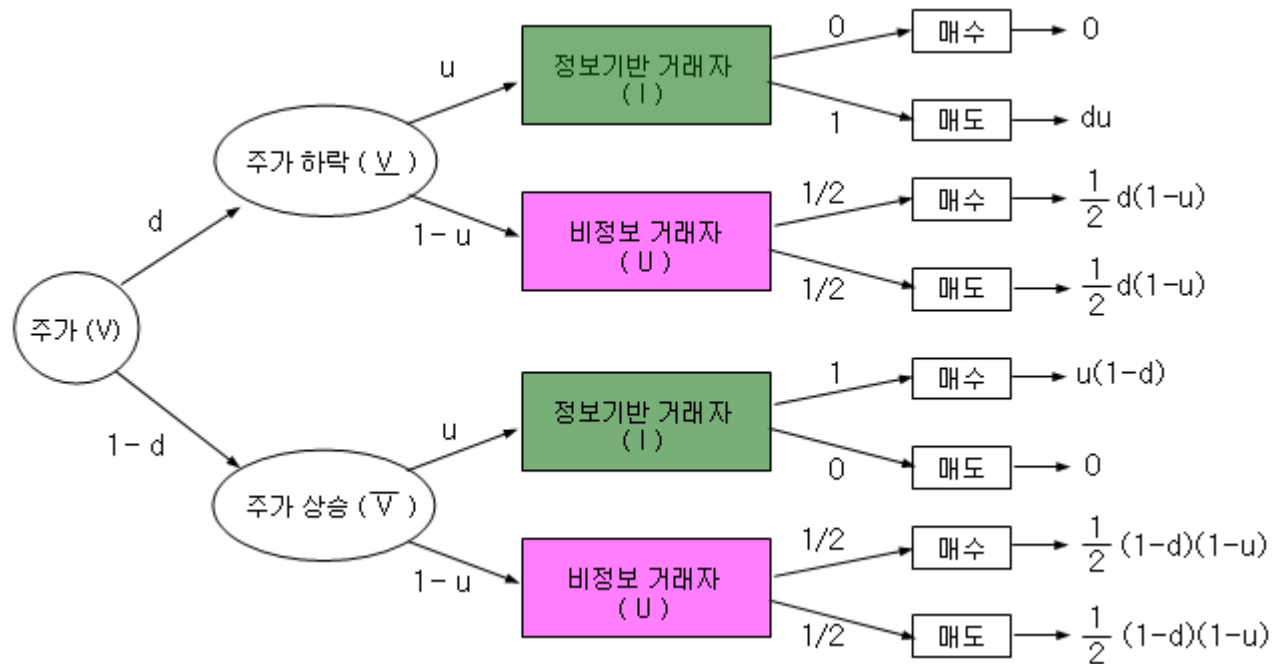
출처 : 조성현님의 블로그 (blog.naver.com/chunjein)

시장미시구조론의 거래자 유형		시장 참여자	특 징	비 고
정보기반 거래자 (Informed Trader)		투자자 (Investor)	<ul style="list-style-type: none"> - 보유한 정보를 이용한 거래 (ex: 내부거래자) - 단기 가격변화에 반응 - 신호매매/ 포지션 매매 - 헤지거래/ 차익거래 - 시장가 주문 성향 - 가격 조성 기능 - 유동성 소비 성향 	<p>참여자 유형 (A) : 거래를 통해 보유 정보가 가격에 반영됨. 보유 정보가 가격에 노출되는 것이 최소가 되도록 전략을 구사.</p>
유동성 거래자 (Liquidity Trader) (Uninformed Trader)	능동적 거래자 (Active Trader)	투자자 (Investor)		<p>참여자 유형 (B) : 정보와는 무관하게 특정 목적을 위한 거래자. (ex : 포트폴리오 자산 배분, 헤지 거래 등)</p>
	공격적 거래자 (Aggressive Trader)	투기적 거래자 (Speculator)		
	수동적 거래자 (Passive Trader)	마켓메이커 (Dealers)	<ul style="list-style-type: none"> - 지정가 주문 성향 - 시장 조성 기능 - 유동성 공급 성향 - 가격 안정 기능 	<p>참여자 유형 (C) : 주가의 가격으로부터 (A)가 보유한 정보를 추출하여 Bid/Ask 가격을 결정하고, 해당 가격에 유동성을 공급하여 스프레드 이득을 취하는 거래자</p>
		투자자 (Investor)		

참고 : Barry Johnson (2010), Algorithmic Trading & DMA (P. 29)

시장 참여자 모형화

출처 : 조성현님의 블로그 (blog.naver.com/chunjein)

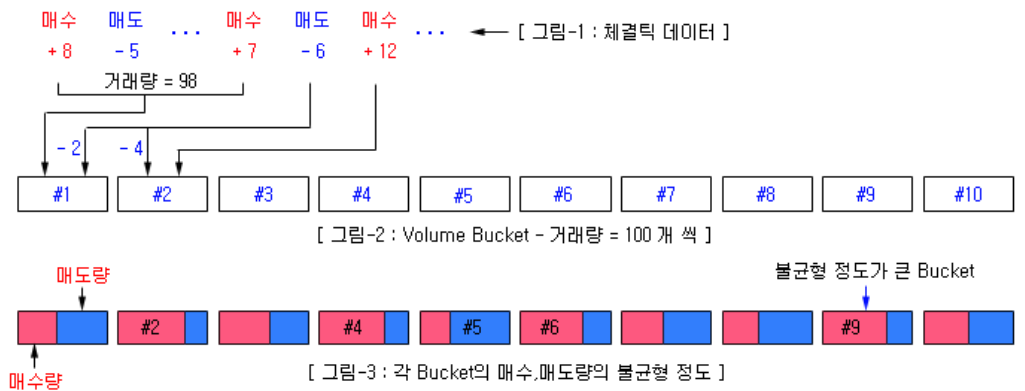


시장 미시 구조 기본

- 기본 상태는 주가 상승과 하락이 50:50
- 정보 기반 거래자가 많을 수록 방향성이 생김
- 주문 비율이나 유동성 등을 분석해서 정보 기반 거래자의 비율 판단
- 이를 이용한 패턴 거래

PIN (Probability of Informed Trading)

- 정보 기반 거래자가 얼마나 존재하는지 추정하는 모델
- 가장 많이 사용하는 것은 VPIN
- $PIN = \text{정보 기반 거래 강도} / \text{전체 거래 강도}$
- 매수 VPIN = volume n 당 매수 / n



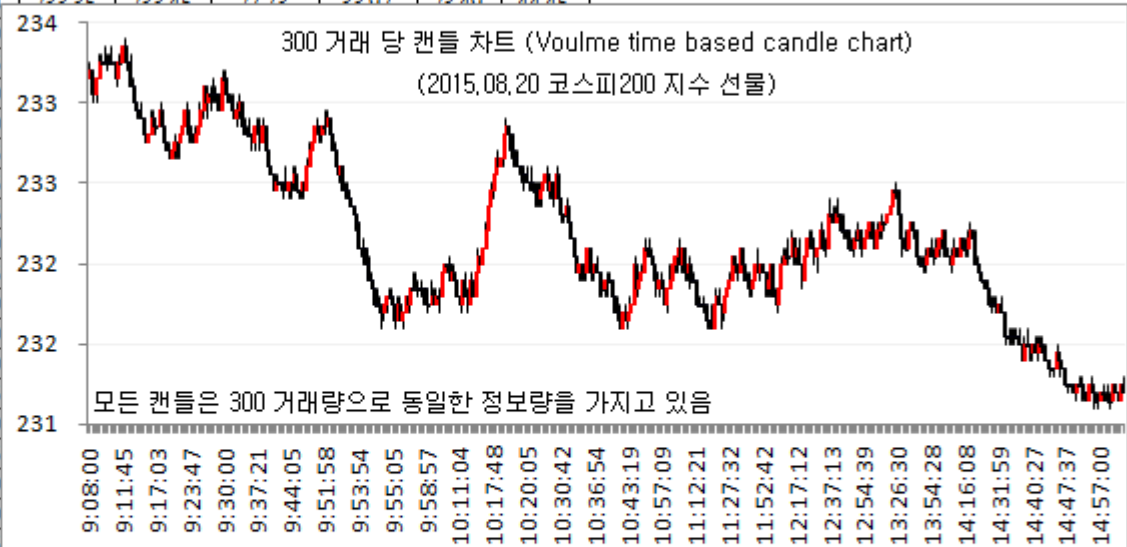
VPIN 실제 데이터

- 300 Volume 당 VPIN 측정
- 2015년 8월 20일 코스피 200 지수
- 평균 34%, 즉 34%의 Informed Trader
- VPIN이 높을 때 주가가 크게 변함

300 거래 당 VPIN 측정 결과

평균	33.97		
표준편차	6.99	경계	1.5

시간	시가	고가	저가	종가	VPIN	평균	Lower	Upper
9:08:00	233.35	233.45	233.35	233.40	28.47	33.97	23.49	44.45
9:08:11	233.40	233.40	233.25	233.30	31.67	33.97	23.49	44.45
9:08:58	233.30	233.35	233.20	233.25	29.47	33.97	23.49	44.45
9:09:19	233.25	233.35	233.20	233.35	27.80	33.97	23.49	44.45
9:09:27	233.35	233.50	233.25	233.45	27.72	33.97	23.49	44.45
9:09:36	233.45	233.50	233.35	233.45	27.72	33.97	23.49	44.45
9:09:47	233.45	233.55	233.35	233.45	27.72	33.97	23.49	44.45
9:10:03	233.45	233.50	233.35	233.45	27.72	33.97	23.49	44.45
9:10:15	233.45	233.50	233.35	233.45	27.72	33.97	23.49	44.45
9:10:41	233.45	233.55	233.35	233.45	27.72	33.97	23.49	44.45
9:10:46	233.45	233.45	233.35	233.45	27.72	33.97	23.49	44.45
9:11:03	233.35	233.45	233.25	233.35	27.72	33.97	23.49	44.45
9:11:24	233.45	233.50	233.35	233.45	27.72	33.97	23.49	44.45
9:11:45	233.45	233.55	233.35	233.45	27.72	33.97	23.49	44.45
9:11:52	233.55	233.60	233.45	233.55	27.72	33.97	23.49	44.45
9:12:10	233.45	233.50	233.35	233.45	27.72	33.97	23.49	44.45
9:13:10	233.45	233.50	233.35	233.45	27.72	33.97	23.49	44.45
9:13:46	233.30	233.35	233.20	233.30	27.72	33.97	23.49	44.45
9:14:03	233.20	233.20	233.10	233.20	27.72	33.97	23.49	44.45
9:14:16	233.15	233.15	233.05	233.15	27.72	33.97	23.49	44.45
9:15:03	233.15	233.20	233.05	233.15	27.72	33.97	23.49	44.45
9:15:06	233.10	233.10	233.00	233.10	27.72	33.97	23.49	44.45
9:15:35	233.00	233.05	232.90	233.00	27.72	33.97	23.49	44.45
9:16:02	232.95	233.00	232.85	232.95	27.72	33.97	23.49	44.45
9:16:09	233.00	233.15	232.90	233.00	27.72	33.97	23.49	44.45
9:16:33	233.10	233.15	233.00	233.10	27.72	33.97	23.49	44.45
9:17:03	233.05	233.10	232.95	233.05	27.72	33.97	23.49	44.45
9:18:13	233.05	233.15	232.95	233.05	27.72	33.97	23.49	44.45
9:18:50	233.15	233.20	233.05	233.15	27.72	33.97	23.49	44.45
9:19:14	233.05	233.10	232.95	233.05	27.72	33.97	23.49	44.45
9:19:34	232.95	232.95	232.85	232.95	27.72	33.97	23.49	44.45
9:19:45	232.90	232.95	232.80	232.90	27.72	33.97	23.49	44.45
9:19:56	232.85	232.90	232.75	232.85	27.72	33.97	23.49	44.45
9:20:20	232.90	233.00	232.80	232.90	27.72	33.97	23.49	44.45
9:21:03	232.95	232.95	232.85	232.95	27.72	33.97	23.49	44.45



실제 전략 – 메릴린치 250번

- 유럽 지수 선물 EURO STOXX 50
 - VPIN을 $n = 2000$ 으로 계산
 - 500개의 VPIN을 선형 회귀로 계산
 - 기울기가 양이고 매수 VPIN이 60 이상일 경우 매수
 - 조건이 깨질 경우 청산
-
- 샵프 지수 3.5



초단타의 꽃 – 마켓메이킹 전략

- 유동성 공급자
- 매수와 매도 주문을 동시에 보냄
- 두 거래가 모두 체결될 경우 스프레드만큼 수익
- 상승장, 하락장 관계 없이 수익
- 재고 관리, 정보 기반 거래자를 방어하는 것이 관건

초단타의 꽃 - 마켓메이킹 전략

Published on TradingView.com, February 15, 2016 10:25 UTC

CME_EOD:EDM2018, W 98.8850 ▼ -0.1250 (-0.13%) O:98.8650 H:99.1650 L:98.7350 C:98.8850



기본 마켓메이킹 전략

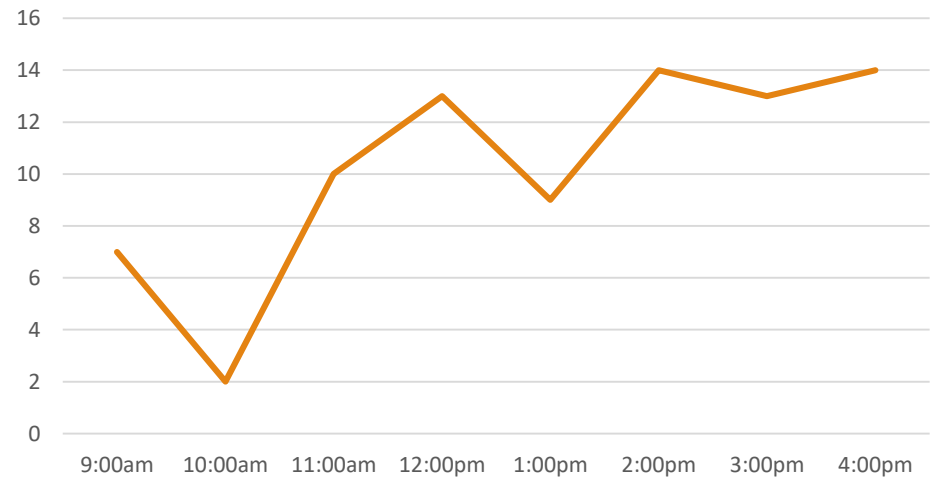
- 가장 최근 거래 가격을 앵커로 지정, 지정된 스프레드만큼 떨어진 가격에 매수 매도를 추가
- 매수나 매도가 이루어진 상태에서 최근 거래 가격이 방향성을 가지고 움직인 경우 앵커 가격을 변경
- 양쪽 거래가 일어난 경우 다시 매수 매도 주문을 추가

기본 마켓메이킹 전략

⇒

건수	수량	가격	수량	건수
5	32	102		
18	89	101		
2	5	100		
		99		
		98	21	5
		97	44	10
		96	182	40

Offer Bid ←



재고 관리 추가

Inventory가 50% 이상 찼을 경우 매수 / 매도 호가 스프레드를 올려줌

시뮬레이션 및 주문 강도를 보고 패러미터를 정함

코스피 선물, 기본 스프레드 2칸, 50% 이상일 경우 3칸으로 시뮬레이션

단방향 재고 관리, 혹은 양방향 재고 관리 모두 가능

VPIN 및 Book pressure 추가

VPIN의 강도 * 2틱 만큼 스프레드에 방어도 추가

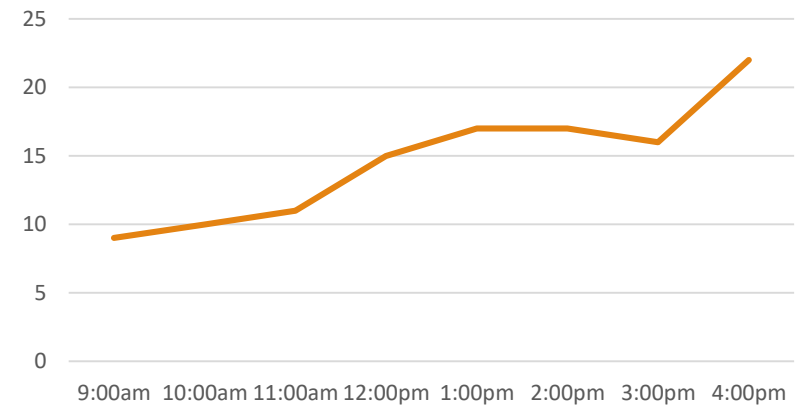
Book pressure가 80% 이상 혹은

20% 미만일 경우 스프레드 1 추가

$BP = \text{TOB Bid volume} / \text{TOB Bid} + \text{Ask volume}$

건수	수량	가격	수량	건수
5	32	102		
18	89	101		
112	1042	100		
		99	13	5
		98	45	22
		97	829	65
		96	157	38

Offer Bid

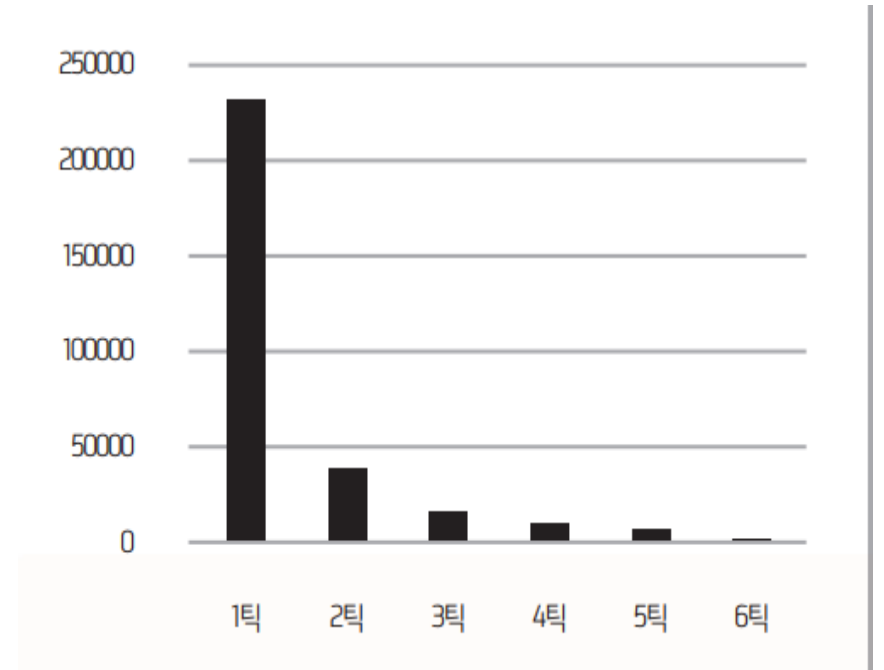


향상된 수익

계단식 마켓 메이킹

과거 통계적 가격 움직임 비율을 통해 스프레드를 단계적으로 지정

추세장이나 변동성이 높은 장에서도 안정적인 수익을 기대

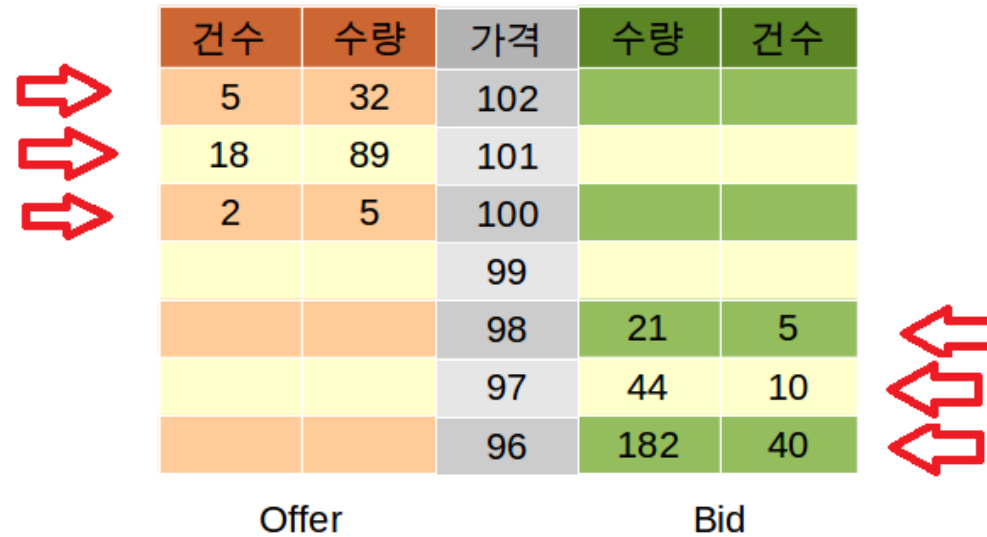


애플 2016년 11월 평균 틱 움직임

계단식 마켓 메이킹

건수	수량	가격	수량	건수
5	32	102		
18	89	101		
2	5	100		
		99		
		98	21	5
		97	44	10
		96	182	40

Offer Bid



Advanced Market Making

- 머신 러닝을 이용한 Order book modeling
- 지정가 매매 / 시장가 매매 강도에 따른 추세
- 정보 기반 거래자의 세분화 (Signal based, True Informed Trading)
- BAB (Bid Ask Bouncing) Rate 계산
- 유동성 측정

추세형 전략

- 시장가 주문 / 지정가 주문 비율에 따라 추세 비율이 다름
- 일반적으로 시장가 주문이 높을 수록 상승 확률이 높음
- 그러나 절대적이지 않으므로 데이터 분석이 필요

다양한 호가의 데이터와 미시 구조론적 팩터를 이용해서 추세 상태를 분석

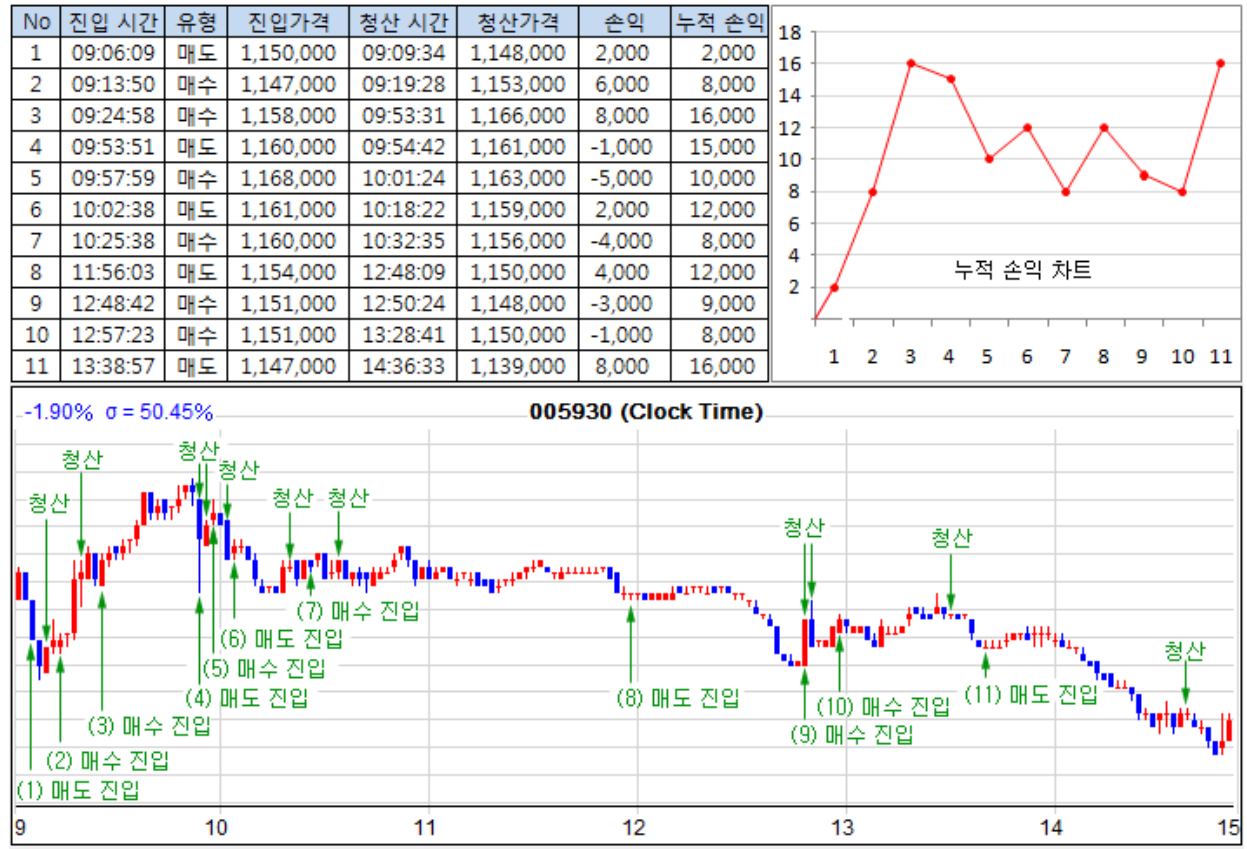
추세 스캘핑 전략

출처 : 조성현님의 블로그

전략 :

- 삼성전자 선물, 최근 200틱 이동 평균이 400틱 이동 평균보다 위 (상승 추세)
- 코스피 200 지수 200틱 이동 평균이 400틱 이동 평균보다 위 (시장의 상승 추세)
- 삼성전자 VPIN이 70% 이상일 경우 매수
- 청산은 Trailing Stop 5틱

추세 스캘핑 전략



2015년 8월 17일 삼성전자 선물

실제 전략 – 이벤트 드리븐 외환 거래

- 외환 거래는 정기적인 뉴스가 있음 (ECO)
- 국가별 발표 5분 이전에 호가와 틱을 체크
- 거대 주문 (> 10 MM USD)이
- + - 5틱 호가 안에 있을 경우 정보 기반 거래자일 확률이 높음
- 같은 방향으로 공격적 주문을 미리 제출

Jan 22, 2018 - Jan 29, 2018				
Date	12:26am	Currency	Impact	
Mon Jan 22	All Day	EUR	📁	Eurogroup Meetings
	6:00am	EUR	📁	German Buba Monthly Report
	8:30am	CAD	📁	Wholesale Sales m/m
	10:14pm	JPY	🔴	BOJ Outlook Report
		JPY	🔴	Monetary Policy Statement
		JPY	📁	BOJ Policy Rate
	11:30pm	JPY	📁	All Industries Activity m/m
Tue Jan 23	▶ 1:30am	JPY	🔴	BOJ Press Conference
	Day 1	All	📁	WEF Annual Meetings
	4:30am	GBP	📁	Public Sector Net Borrowing
	5:00am	EUR	📁	German ZEW Economic Sentiment
		EUR	📁	ZEW Economic Sentiment
	All Day	EUR	📁	ECOFIN Meetings
	6:00am	GBP	📁	CBI Industrial Order Expectations
	10:00am	EUR	📁	Consumer Confidence
		USD	📁	Richmond Manufacturing Index
	6:30pm	AUD	📁	MI Leading Index m/m
6:50pm	JPY	📁	Trade Balance	

기술적 분석 vs 미시 구조론

시장 미시 구조론

- 주문 흐름
- 호가 데이터
- 시장 참여자

기술적 분석

- 패턴
- 이동 평균선
- 거래 지표
- 지지선, 저항선

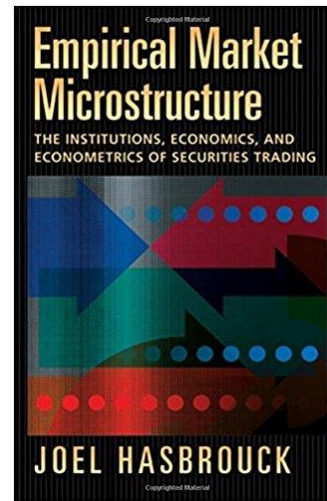
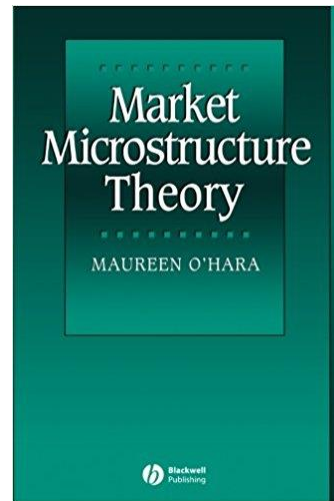
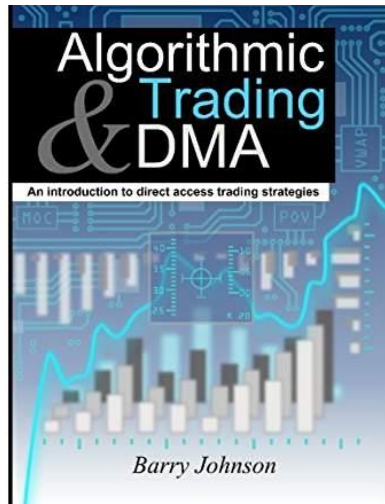
시장 미시 구조 정리

- 가격이 형성 되는 과정을 연구하여서 수익을 얻는 전략
- 틱 데이터 / 주문 데이터 / 호가 데이터를 이용하여 분석
- 수년전부터 동양증권 등 증권사가 틱 데이터를 제공하거나 판매하기 시작
- 국내에서는 높은 수수료와 세금으로 인해 HFT로 수익을 얻기 쉽지 않음
- 그러나 미시 구조에 대한 연구로 비용을 줄이거나 시장 상황을 파악할 수 있음
- HFT가 아니더라도 Mid frequency Level의
기술적 분석 베이스 전략을 연구를 통해 검증하거나 구성할 수 있음

시장 미시 구조는 방대하다

조성현님의 블로그 (blog.naver.com/chunjein)

미시구조 관련 책



데이터 분석을 이용한 전략 수립

아이디어 수집

-논문

- Time horizon 바꾸기
- Instrument 바꾸기
- Market 바꾸기
- Parameter 바꾸기

-Economical 아이디어

- 뉴스 / 정책
- Correlation / Pair Trading
- Fundamental / Factor Driven Strategy

데이터 분석을 이용한 전략 수립

-Heuristic / Post trade analysis

- 경험적 패턴
- 손실 거래 / 슬리피지 분석

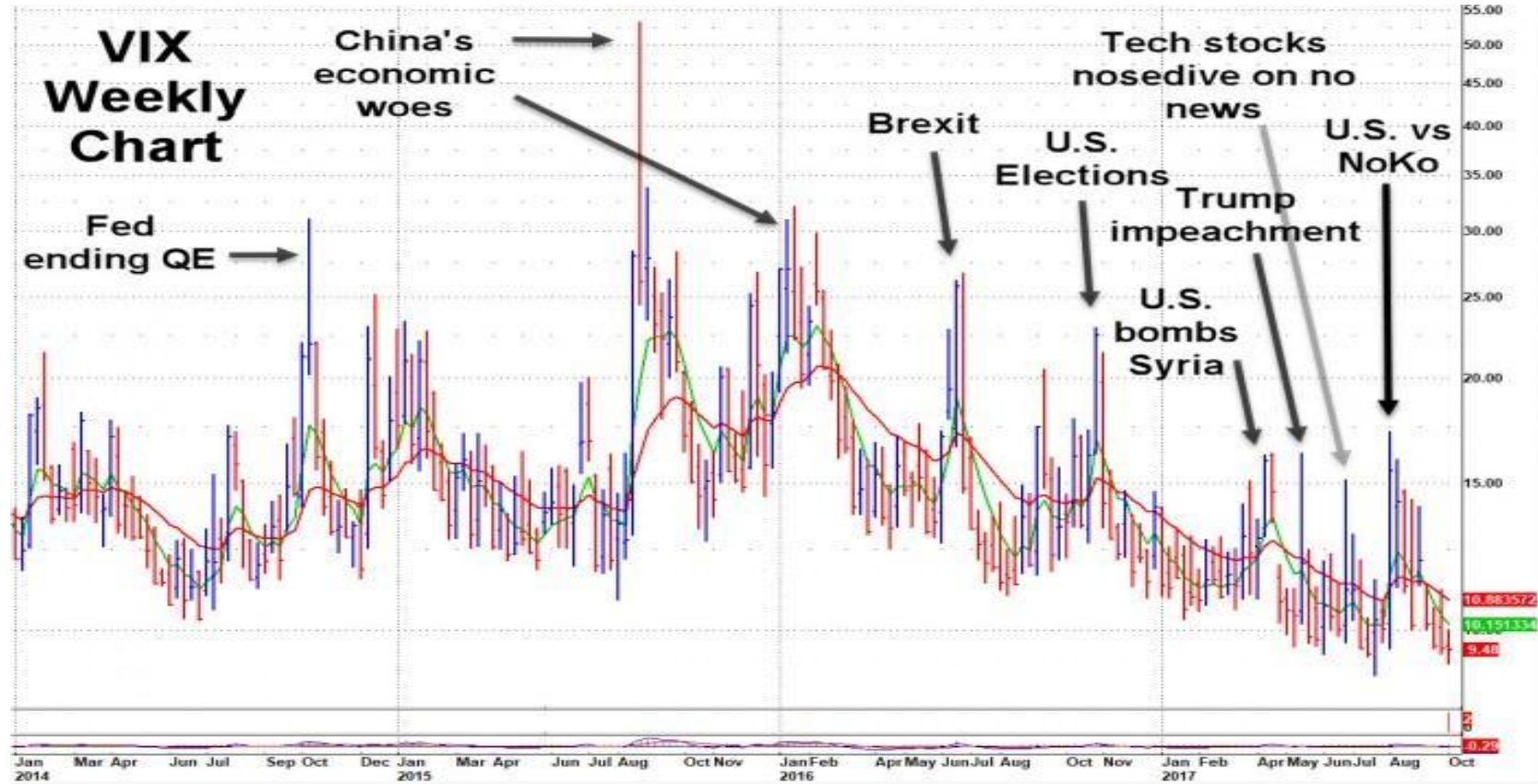
-머신러닝 / AI 툴 이용

- 성공 확률이 낮음
- 앞선 거래 전략 수립 방식의 검증 / 파라미터 최적화 용도로 사용

실제 전략 – VIX 매도 전략

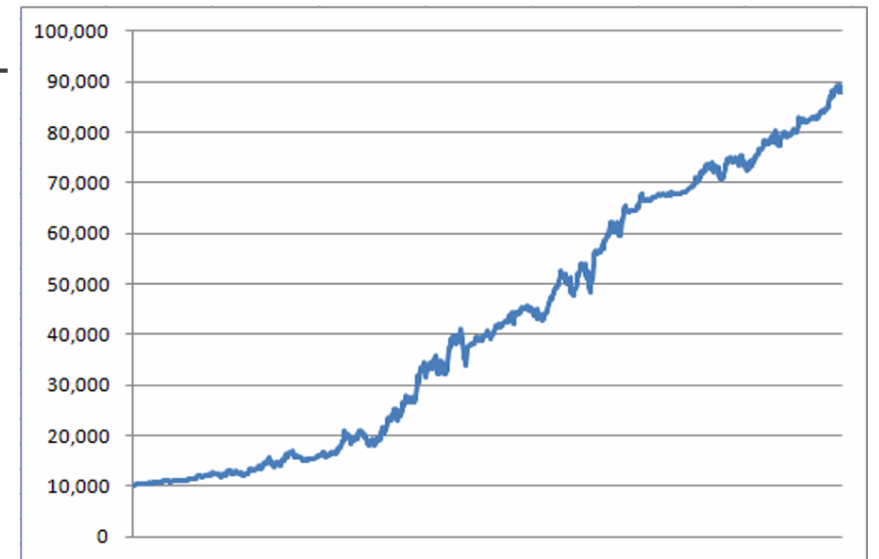
- VIX는 안정화된 시간에서 우하향
- 그러나 Volatility가 생길 때 마다 Volatility spike가 일어나면서 상승
- 3일 연속 상승세, 혹은 뉴스 근처일 경우 청산, 이외에는 매도 상태를 유지
- Paper : VIX Trading – Credit Suisse 등 수많은 논문에서 검증된 방식

실제 전략 – VIX 매도 전략



실제 전략 – VIX 매도 전략

- 이 논문의 전략의 Time horizon을 분당으로 바꿔봄
- Parameter를 이동평균선과 VPIN으로 바꿈
- 뉴스 Detection을 Twitter Toxicity로 바꿔봄

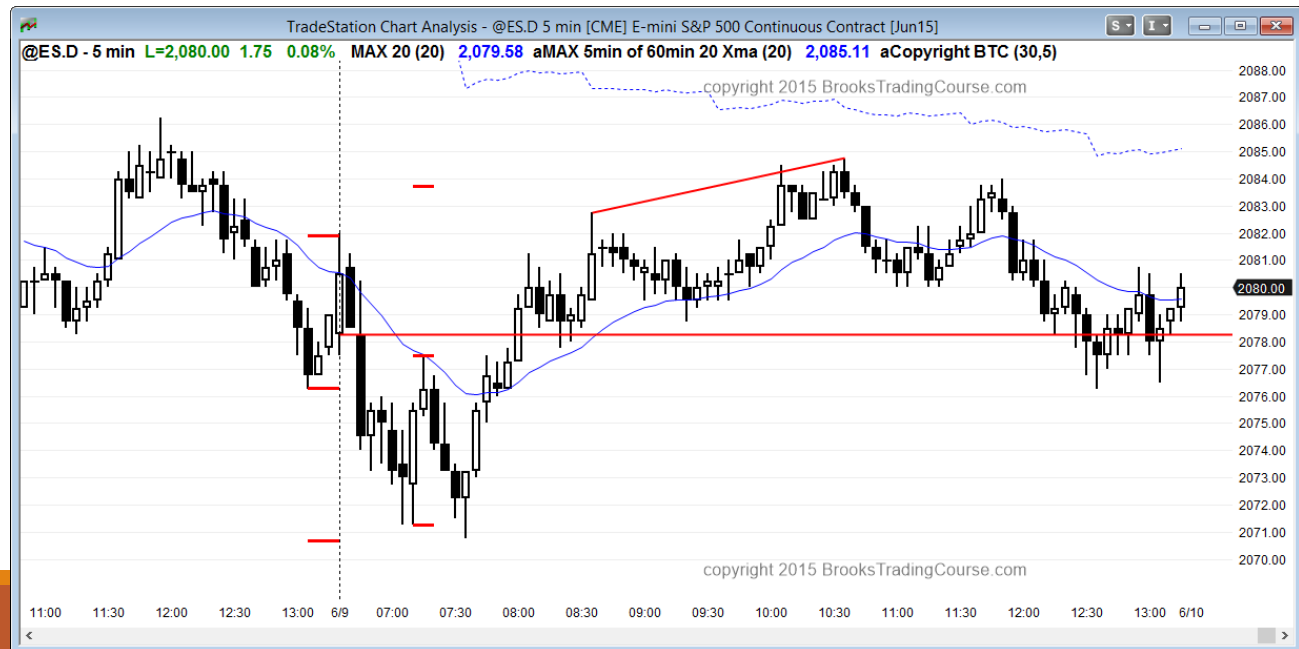


2014년 2월 – 3월까지 Dollar Value 초단타 VIX 전략

실제 전략 - 머신러닝을 이용한 Pattern Recognition

- 가정 : 시간대별 E-mini 선물 거래량에 패턴이 있다
- Feature : 시간, Order Imbalance (~ VPIN), Normalized 가격
- CNN 인공 신경망 알고리즘 사용,
장 개시와 점심 시간 근처의 패턴을 잡아냄

- 샤프 지수 약 2.3



전략 검증을 위한 데이터 수집

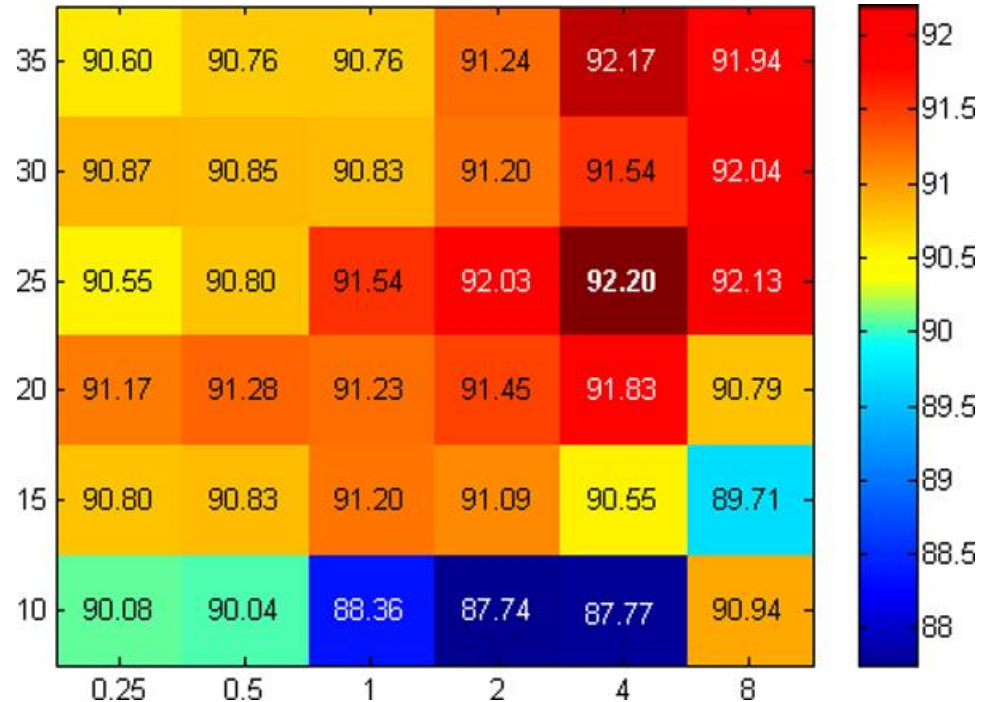
- 일반인이 가장 어려움을 겪는 단계
- 체계적 거래를 위해선 기록, 그러나 테스트를 위해선 구입을 추천
- 분봉 데이터까지는 쉽게 구할 수 있으므로 틱 데이터 중심의 전략 보다는 기술적 분석을 기반으로 한 전략 검증이 접근 용이

주성분 분석

- 수많은 데이터를 한번에 분석하기엔 시간 / 공간 / 비용이 많이 든다
- 연관성이 높은 데이터들을 줄여서 차원을 줄이는 단계가 필요
- PCA Analysis (주성분 분석)은 널리 알려진 방법
- 예 : 수학 성적, 과학 성적, 키, 인종 등의 관계를 구하려 할 때, 수학 성적과 과학 성적은 상관관계가 높으므로 줄여줄 수 있음
- 머신러닝 등에서 많이 이용

최적화 및 백테스트 과정

- Heatmap style 최적화 방식을 추천 – python matplotlib에서 자동 제공



전략 검증 / Robustness

- Overfitting

- 퀀트 트레이딩 / 데이터 분석에서 가장 주의해야할 적
- 데이터에 지나치게 맞춰서 파라미터를 설정하여 시뮬레이션 상으로 고수익인 경우
- In sample / out sample 및 Boot strapping 적극 이용

- Data mining

- 우연히 데이터 상으로 관계가 생길 수 있음
- Economical Reasoning 과정을 거쳐야 함
- 통계적 요소를 체크

이제 데이터 분석은 필수

데이터의 대중화로 간단한 분석이 기본이 되는 시대
Naïve Decision Making으로 투자에 성공하기 어려워짐

기본 분석 툴에 익숙해져야 한다.
데이터 과학자의 마인드를 가져야 한다.

Thank you
Questions?