

- NIPS  
- KDD

# 机器学习与量化交易实战

第九课

- AAAI 2017

- ICML Dec.

- IJCAI

— ~~AA~~ 特征值

SVD

10

—  $w_i$

— 5 paper

# 特征值 & 特征向量

$$A_{100 \times 100} \vec{x}_{100 \times 1} = \vec{u}_{100 \times 1}$$

~~$f: \vec{x} \rightarrow \vec{u}$~~

$\rightarrow A_{\mathbb{R}^{100 \times 100}}$

mathematical object

$a$   $a^*$

$$A_{100 \times 100}$$

$$a(a+ib)^{-1} \mathbb{I}$$

$$A \vec{x}_{100 \times 1} = \lambda \vec{x}_{100 \times 1}$$

$$\vec{x} \parallel A \vec{x}$$

$$A \vec{x}$$

$$(A - \lambda I) \vec{x} = 0$$

$$A \vec{x} = \lambda \vec{x}$$

$$A \vec{x} - \lambda \vec{x} = 0$$

$$(A - \lambda I) \vec{x} = 0$$

$$(A - 6I) x = 0$$

$$Ax = 2x$$

$$\det(A - \lambda I) = 0$$

$$\begin{vmatrix} 1-\lambda & 2 & 3 \\ 5 & 6-\lambda & 7 \\ 9 & 9 & 9-\lambda \end{vmatrix} = 0$$

$$1 \ 2 \ 2$$

$$\lambda : n$$

$$\lambda = 6, 6, 2$$

$$3 \times 3 \quad Ax = \lambda x$$

$$1 \ 2 \ 6 \quad Ax = 1x$$

$$\begin{vmatrix} 1 & 2 \\ 2 & 2 \end{vmatrix}$$

$$| A = A^T | \quad \textcircled{1} \quad \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}^T \neq \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

$$\lambda \in \mathbb{R}$$

$$\lambda_i \rightarrow \vec{x}_i \text{ 互相垂直的}$$

对于非对称  $A_{n \times n}$

$$| A = S \Lambda S^{-1} |$$

$$\rightarrow [\vec{x}_1, \vec{x}_2, \dots] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_n \end{bmatrix} S^{-1}$$

$$\boxed{Ax = \lambda x} \quad A = S \Lambda S^{-1}$$

$$AS = A \cdot [\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n]$$

$$= [\lambda_1 \vec{x}_1, \lambda_2 \vec{x}_2, \dots, \lambda_n \vec{x}_n]$$

矩阵  
乘法  
结合

$$\downarrow$$

$$= [\vec{x}_1, \vec{x}_2, \dots] \cdot \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_n \end{bmatrix}$$

$$= S \Lambda \Rightarrow \boxed{AS = S \Lambda}$$

$$\boxed{A = S \Lambda S^{-1}}$$

$$\boxed{A = A^T} \quad \boxed{A = S \Lambda S^{-1}}$$

$$S^{-1}$$

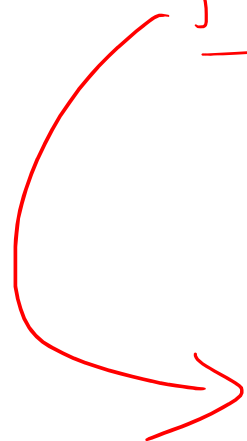
S: 正交阵

$$S^{-1} = S^T$$

$$\boxed{A = S \Lambda S^{-1}}$$

$$\boxed{A = Q \Lambda Q^T}$$

$$\vec{v}_i \text{ 是 } \vec{v}_i \text{ 的 } \vec{v}_i$$



$$\vec{x}_i$$

$$\lambda$$

$$A^k$$

$$A^2$$

$$= A \cdot A$$

$$= S \Lambda S^{-1} S \Lambda S^{-1}$$

$$\boxed{A = S \Lambda S^{-1}}$$

$$A^2 = S \Lambda^2 S^{-1}$$



MIT 公开课. LA

SVD 任意矩阵

$$A = A^T$$

$$A = Q \Lambda Q^T$$

$$\Rightarrow A = U \Sigma V^T$$

正交阵

正交阵

对称

~~$Ax = xx$~~

$$U, \Sigma, V = \text{SVD}(M)$$

$\vec{u}_1, \vec{u}_2, \vec{u}_3$   
 $R^n$

(A)

$R^n$

$A\vec{u}$

$$A = U \Sigma V^T$$

$$Ax = \lambda x$$

$\vec{u}_i$

$$A\vec{u}_i = \sigma_i \vec{v}_i$$

$\vec{v}_i$

$\perp \vec{u}_i$

$A^T A$

$A A^T$

9 6 3 1 1 1 0.3 0.1

$\Delta$

$$A = U \Sigma V^T$$

$\vec{u}_2$   
 $\vec{u}_3$   
 $X$

[3 6 9 2 5]

$\updownarrow$

$\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n$

[1 2 3 4 5]

$$AU = A[\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n]$$

$$AU = V\bar{\Sigma}$$

$$\begin{aligned} &= [\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n] \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \\ & & & \sigma_n \end{bmatrix} \begin{array}{l} Au_1 = \sigma_1 v_1 \\ Au_2 = \sigma_2 v_2 \\ \vdots \end{array} \\ \textcircled{A^T A} = \textcircled{B} \quad B^T = A^T A = \textcircled{Q \Lambda Q^T} \end{aligned}$$

$$A[\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n] = [v_1, v_2, \dots, v_n] \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}$$

$$AU = V\bar{\Sigma}$$

$$\textcircled{A^T A}$$

$$A \wedge$$

$$\textcircled{4}$$

$$\underline{A} = V\bar{\Sigma}U^{-1} = V\bar{\Sigma}U^T$$

Trick

$$\underline{A^T} = U\bar{\Sigma}V^T$$

$$AA^T = V\bar{\Sigma}^2V$$

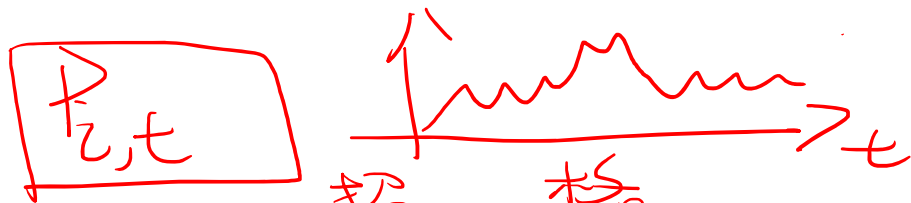
$$\textcircled{A^T A} = U\bar{\Sigma}V^T \cdot V\bar{\Sigma}U^T = \textcircled{U\bar{\Sigma}^2U^T}$$

MIT 线性代数

Portfolio

(N) (20)

Stock - 2



$$r = \frac{P_t - P_{t-1}}{P_{t-1}}$$

$t: [P_{1,t} \quad P_{2,t} \quad P_{3,t} \dots P_{N,t}] \rightsquigarrow$

$[r_{1,t} \quad r_{2,t} \dots r_{N,t}]$

$\vec{u} = [\ ]_N$

[ \_\_\_\_\_ ]

$\sum [\ ]_{N \times N}$

[ \_\_\_\_\_ ]

平均

[ \_\_\_\_\_ ]

$\downarrow M$

$t_1 \quad t_2 \quad t_3$   
2 6 7  
9 8 10  
15 20

mean  
var

$\{t, [r_1, r_2, \dots, r_n]\} \rightarrow$  十六进制矩阵

$t, [w_1, w_2, \dots, w_n] = \vec{w}_{n \times 1}, \lambda: 98765$



$$\begin{array}{c}
 \begin{pmatrix} \bar{1} \\ \bar{2} \end{pmatrix} \\
 \bar{u}
 \end{array}
 \begin{array}{c}
 \begin{matrix} 201511 \\ t_1 \end{matrix} \\
 t_2 \\
 \vdots \\
 M
 \end{array}
 \begin{array}{c}
 \begin{bmatrix} r & r & r & r \end{bmatrix} \\
 \begin{bmatrix} \end{bmatrix} \\
 \vdots \\
 \begin{bmatrix} \end{bmatrix}
 \end{array}
 \end{array}$$

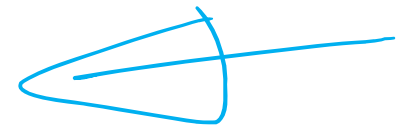
$$5 \quad 5 \quad 6 \quad 7 \quad 9 \quad \rightarrow \quad \bar{x} \quad \text{var}$$

$$\Sigma_{n \times n} = \Sigma^T$$

$$\text{Risk} = w^T \bar{\Sigma} w$$

Stanford  
convex  
optimization

$$\begin{aligned} \min_w \quad & w^T \bar{\Sigma} w \\ \text{s.t.} \quad & w_1 + \dots + w_n = 1 \end{aligned}$$



$$\vec{w}^* = \frac{\bar{\Sigma}^{-1} \cdot \mathbf{1}}{(\bar{\Sigma}^{-1} \cdot \mathbf{1})^T \cdot \mathbf{1}}$$

~~Rank~~

Random  
matrix  
~~theor~~  
theory

$$\rho(\lambda) = \begin{bmatrix} \phantom{\lambda} \end{bmatrix}$$

$$\left\{ \begin{array}{l} 0.1 \\ 0.25 \end{array} \right\}$$



pdf

$$\Sigma = \lambda$$

$$\Sigma = Q \Lambda Q^T$$

$$\Sigma' \neq$$



~~Marchenko~~

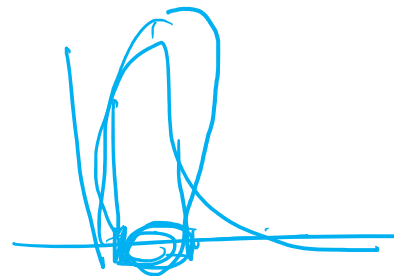
Marchenko-Pastur



$$\Sigma = \Sigma^T$$

1)

$$\text{SVD}(\Sigma)$$



$$\Sigma = Q \Sigma Q^T$$

$$\lambda_1 \dots \lambda_n$$

100

$$\Sigma'$$

$$2) \quad 60-70 \quad \lambda_1 \quad \lambda_2 \quad \lambda_5 \dots \lambda_{80}$$

$$\lambda_i = 0 \quad \lambda_i \in [\quad]$$

$$3) \quad \Sigma' \rightarrow$$

$$\underline{\sigma_1 = 2}$$

$$\underline{\sigma_2 = 2.5}$$

$$\underline{\vec{x}_1 = [1, 1]}$$

$$\underline{\vec{x}_2 = [-1, 1]}$$

$$A = A^T$$

$=$

$$A = Q \Lambda Q^T$$

$\sigma$

$x$

$$= \begin{bmatrix} 1 & \textcircled{1} \\ 0 & \textcircled{0} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2.5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\uparrow$$
  

$$\vec{x}_1$$

$$\uparrow$$
  

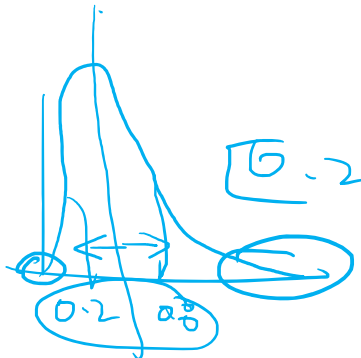
$$\vec{x}_2$$

$=$

①  $\Sigma$  covariance matrix

②

③



[0.2 0.8]

$\sigma_1$   
 $x_1$

$\sigma_2 \dots \sigma_n$   
 $x_2 \dots x_n$

$\sigma_1$	0	0	0	0	$\sigma_2$	...
x	x	.	.	.	x	

④

$\Sigma'$

$Q \Lambda Q^T$

$w$

⑤

$\Sigma'$

$\rightarrow$

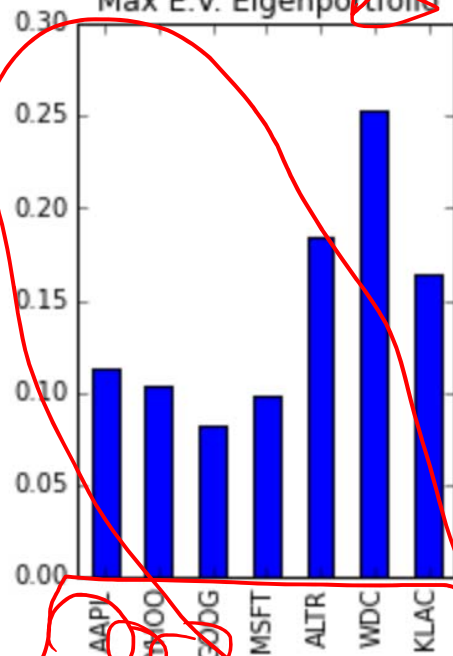
$risk = w^T \Sigma' w$

$$\lambda \rightarrow \bar{u} \rightarrow \bar{w}_{0.2}$$

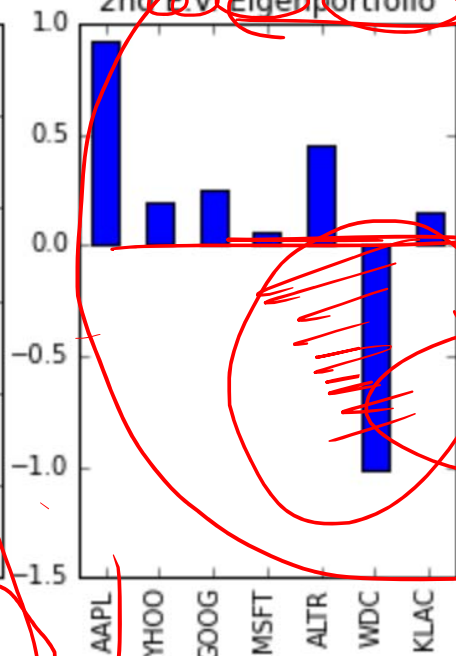
0.3  
0.4

$$\begin{bmatrix} 3 \\ 0.3 \\ 0.3 \end{bmatrix}$$

Max E.V. Eigenportfolio



2nd E.V. Eigenportfolio



Short

$$\begin{bmatrix} 0.5 \\ 3 \\ 6 \\ 6.5 \\ 3 \end{bmatrix}$$

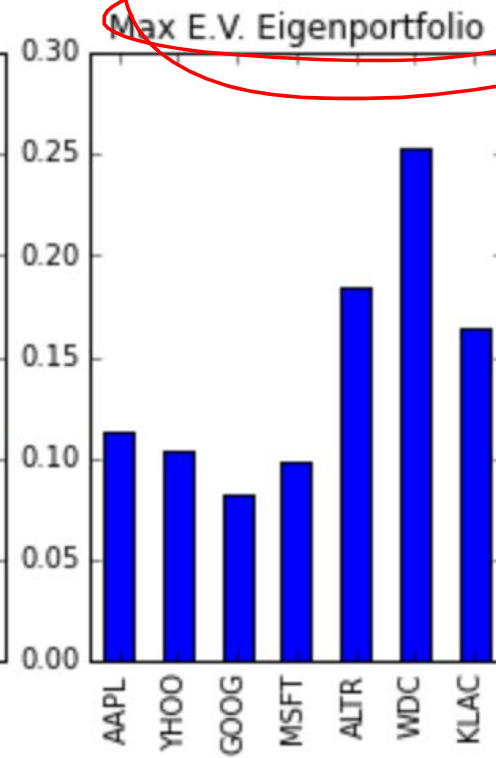
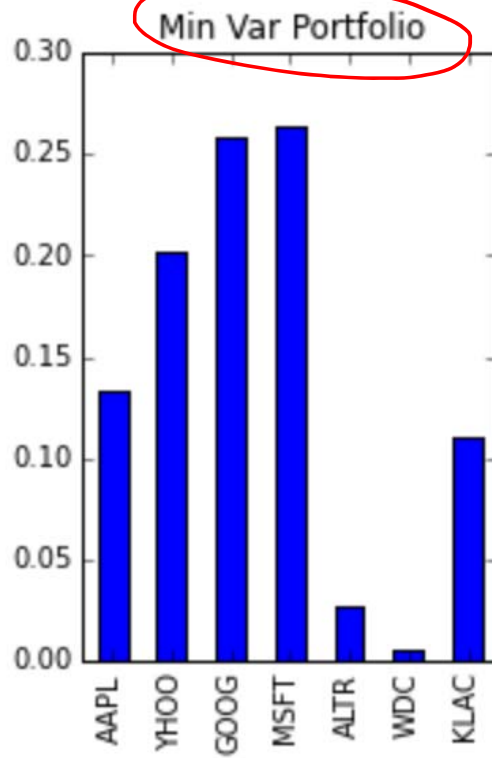
$$\lambda: 0.2, 1.1$$

$$\vec{x} = \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix} \begin{bmatrix} 0.1 \\ 0.9 \end{bmatrix}$$

sell

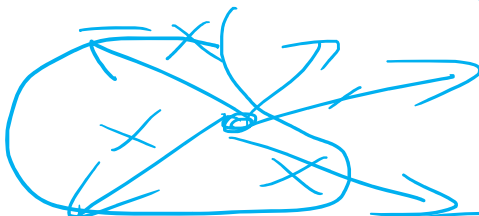
Min Var

Einsen

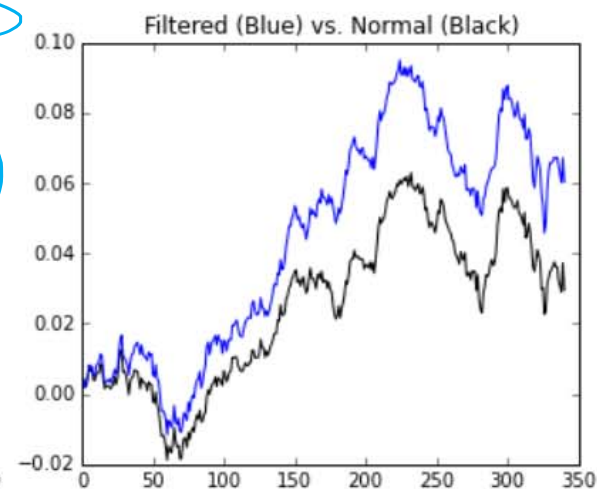
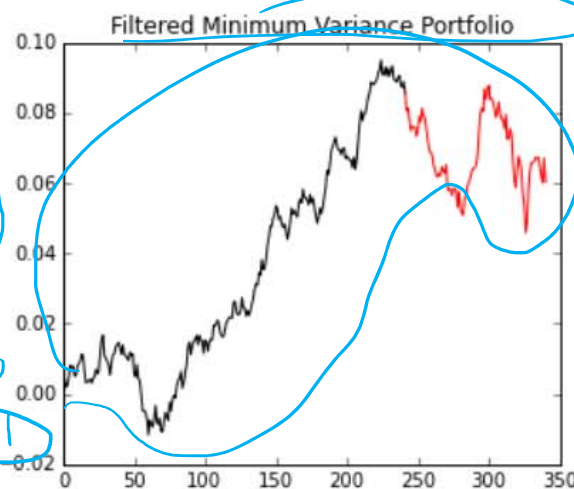
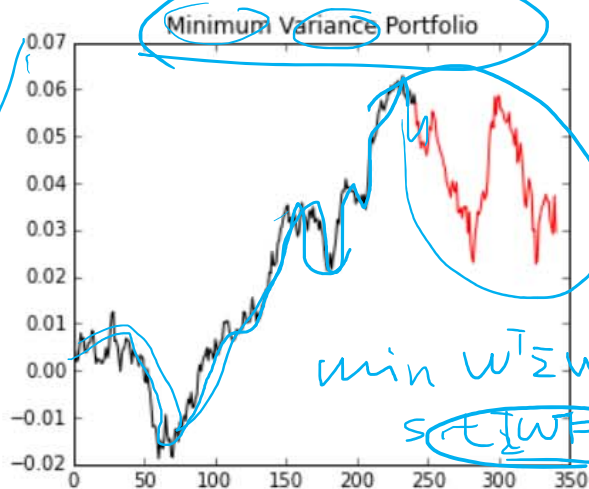


$$A' = w$$

$$A = U \Sigma U^T$$



$$w \Sigma$$



Subject  
to

$$A' = Q \Lambda Q^T$$

1603.08604v1.pdf

Symbol	Futures	Annualized Sharpe Ratio	Capability Ratio
PL	Platinum	3.29 (1.12)	12.51 (4.27)
CNQ	E-mini NASDAQ 100 Futures Quotes	2.07 (2.11)	7.89 (8.03)
AD	Australian Dollar	1.48 (1.09)	5.63 (4.13)
BP	British Pound	1.29 (0.90)	4.90 (3.44)
ES	E-mini S&P 500 Futures Quotes	1.11 (1.69)	4.22 (6.42)

KNL

- The overall training dataset consists of the aggregate of feature training sets for each of the symbols. The training set of each symbol consists of price differences and engineered features including lagged prices differences from 1 to 100, moving price averages with window sizes from 5 to 100, and pair-wise correlations between the returns and the returns of all other symbols. The overall training set contains 9895 features. The motivation for including these features in the model is to capture memory in the historical data and co-movements between symbols



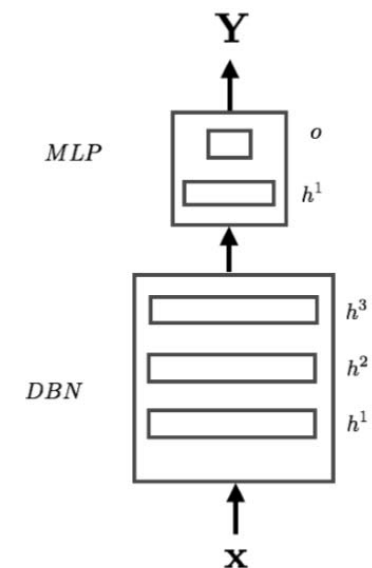
150612a.pdf

## Deep Learning for Multivariate Financial Time Series

$$\mathbf{x} = \left\{ r_d^i(t), \dots, r_d^i(t-19), r_m^i(t-2), \dots, r_m^i(t-13), I(t=J) \right\},$$

Model	Neurons	$\epsilon_p$	$\epsilon_f$	$E_V$ (%)	$E_T$ (%)
Naive	-	-	-	50.03	50.93
LReg	-	-	-	49.96	50.74
MLP	20	-	0.1	49.66	50.84
DBN	400	$10^{-11}$	$10^{-3}$	46.52	47.11

Table 5.5: *Result from DBN-MLP and benchmarks.*





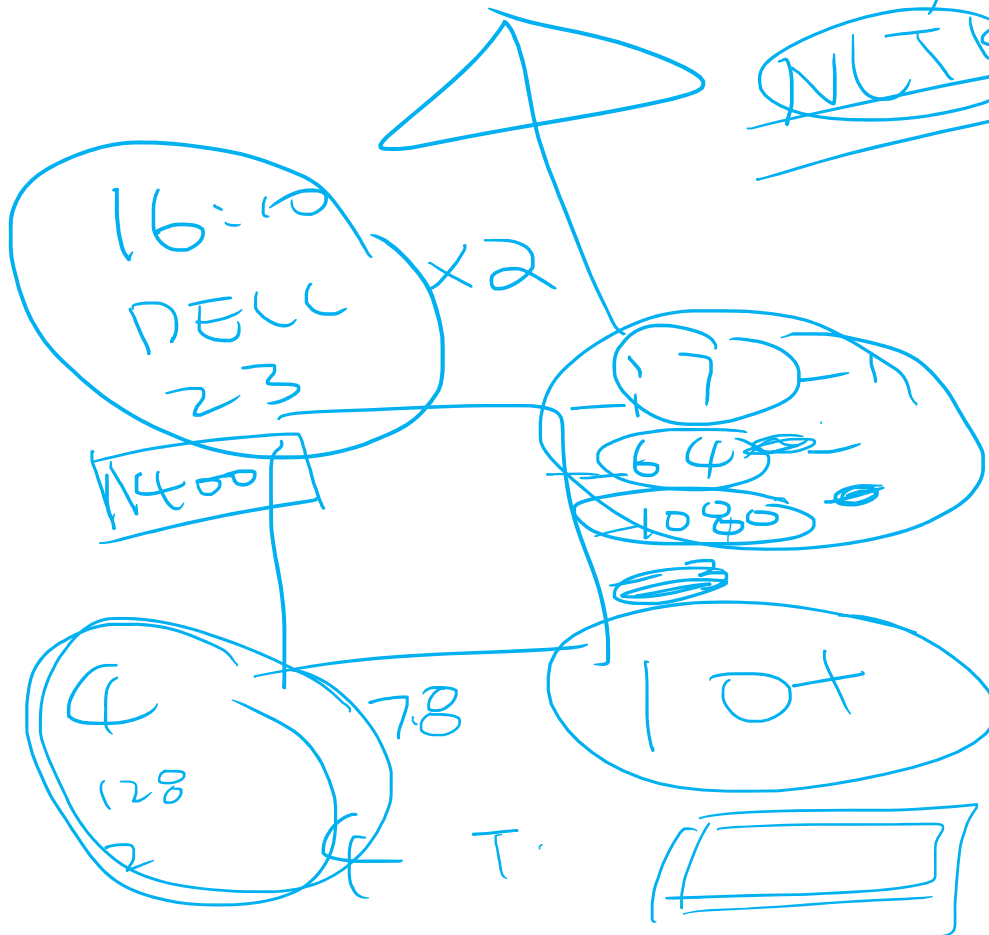
ecbsp9.en.pdf

Keras  
Conclusion

gensim

NLTK

The reliability and accuracy of existing computational measures of investor sentiment leaves much to be desired. We therefore propose a direct and unambiguous measure of investor sentiment, namely the relative frequency of occurrence of two terms commonly used by investors in Twitter updates and Google search queries. Daily Twitter bullishness is indeed found to be a useful investor sentiment indicator. Our analysis shows a positive correlation between Twitter bullishness and Google bullishness on a weekly basis; furthermore, it finds that the former leads changes in the latter. In addition, the two indicators of bullishness from different data sources are found to be positively correlated to existing surveys of investor sentiment, such as the Daily Sentiment Index and the US Advisors' Sentiment Report of Investors Intelligence. More importantly, we find that daily Twitter bullishness leads stock index returns in the United States (Dow Jones, SP500, Russell 1000, Russell 2000), the United Kingdom (FTSE 100) and Canada (GPSTSE), but has only very modest predictive value in respect of the Chinese stock market, as expected. While high Twitter bullishness predicts an increase in stock returns, we observe that these return to fundamental values within a week. Our research thus appears to support the hypothesised role of "investor sentiment" in behavioural finance. We also note the strong positive linear correlation between Google bullishness and Chinese stock index prices ( $\gamma = 0.65$ ,  $p = 0.01$ ), with the former apparently leading the latter in extreme market conditions. This result suggests the merits of studying the predictive value of online information sources such as Weibo for the Chinese market; a topic that has received little interest in the literature.



fora\_gyula\_krisztian.pdf

Indicator Name	Formula
EMA(N)	$EMA_t(N) = \frac{1}{N+1} \sum_{i=0}^{\infty} p_{(t-i)} \left(1 - \frac{1}{N+1}\right)^i$
MACD(N,M)	$MACD_t(N, M) = EMA_t(N) - EMA_t(M)$
K%	$\%K(N) = \frac{100(c_t - \max_{i=0..N}(p_t))}{(\max_{i=0..N}(p_t) - \min_{i=0..N}(p_t))}$
Arms Index	$\frac{NR_{Advances} \times VOL_{down}}{NR_{Declines} \times VOL_{up}}$

Research QVANT

