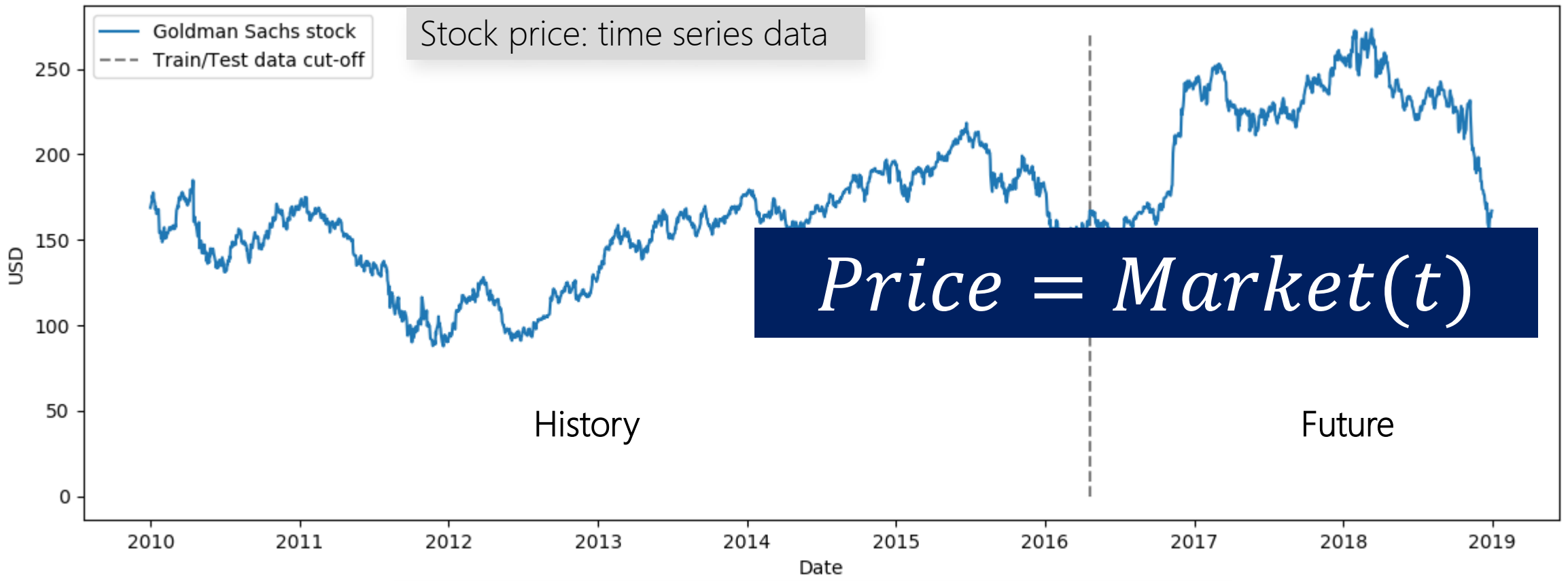



What You Say and How You Say It Matters: Predicting Stock Volatility Using Verbal and Vocal Cues

ACL 2019, Poster, by Yu Qin(info, RUC), Yi Yang*(business, HKUST)

Cai Mingshi / i@unoiou.com / 2019/11/12

Stock Prediction Task



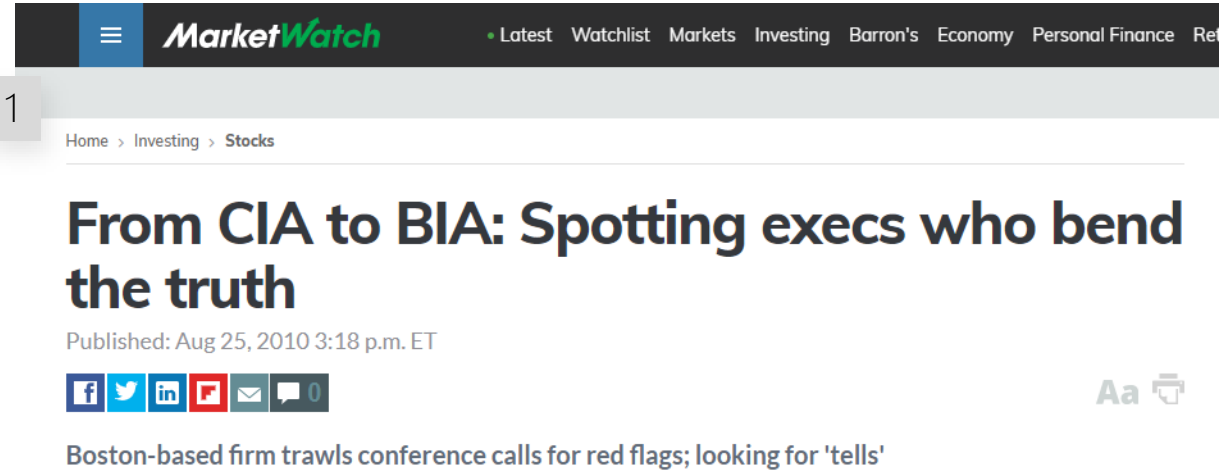
Goldman Sachs(GS) stock close price during 2010-2019

Contents

Ideas

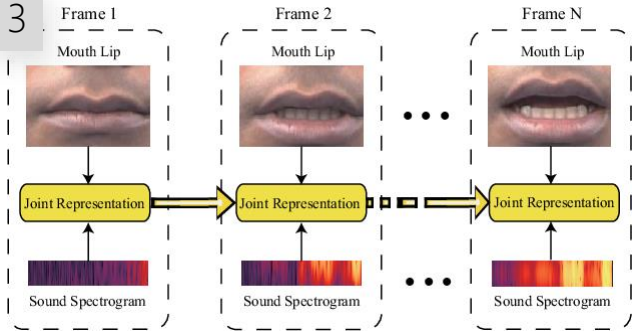
From CIA to BIA, Integrating Audio Features

1



2

3



Alphabet 2019 Q3

Baidu 百度 2019 Q3

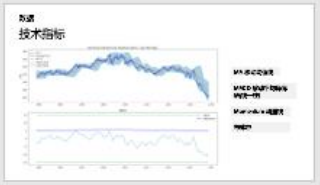
2019 Q3

Related Works

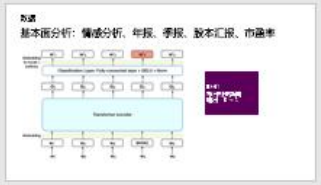
Empirical Methods

1 ★

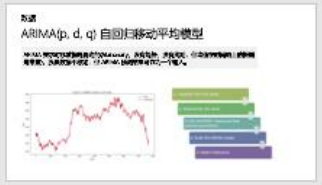
2 ★



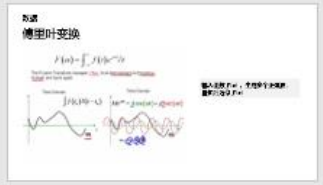
3 ★



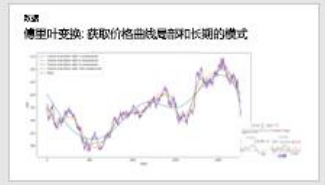
4 ★



5 ★

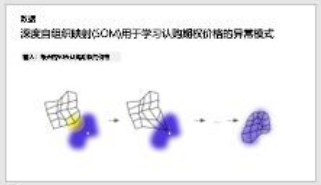


6 ★

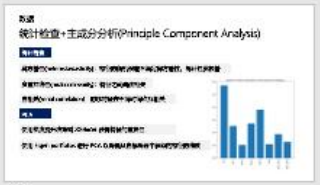


7 ★

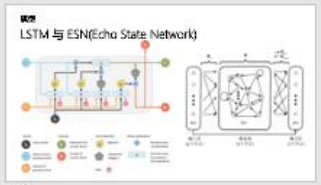
8 ★



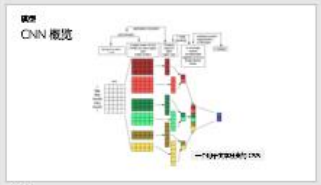
9 ★



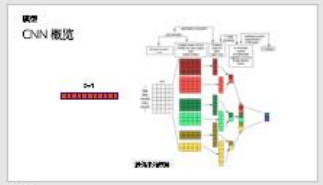
10 ★



11 ★

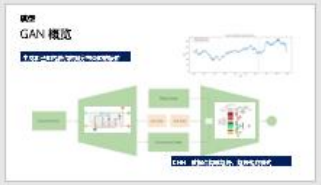


12 ★



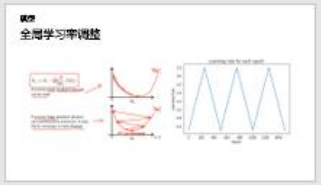
13 ★

14 ★



15 ★

16 ★



17 ★

18 ★

19 ★


20 ★

21 ★

Statistical Modals

Classical Machine Learning

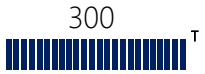
Financial

Deep Learning

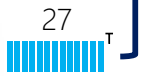
Data & Data Process

S&P 500, 2017, 4Q

Earnings call transcript → Sentence level → Pre-trained GloVe-300



Earnings call audio → Audio clips by CEO → Praat(pitch, amplitude, intensity, etc.)



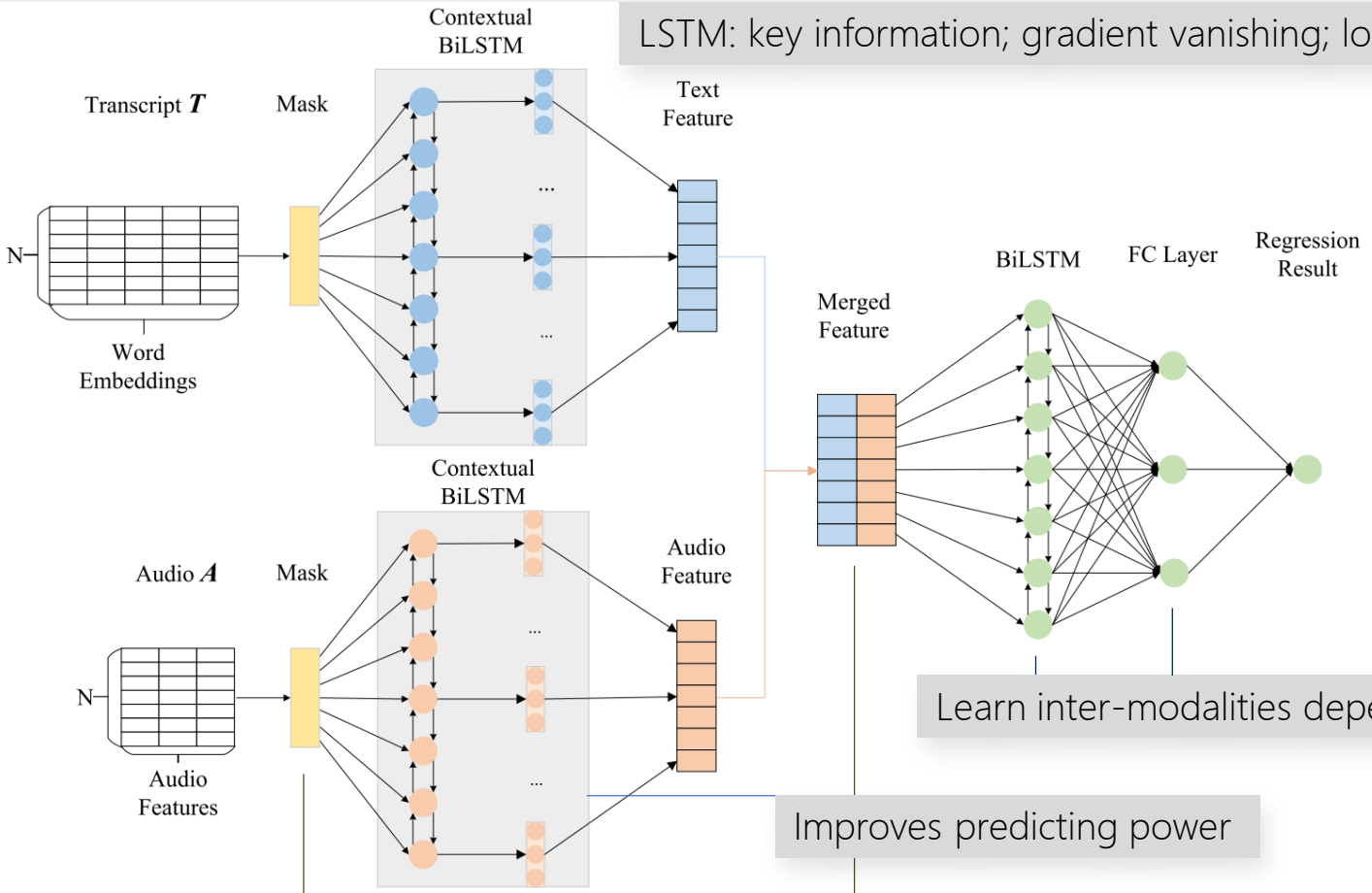
Align (IFA)

History stock price → $v_{[t-\tau,t]} = \ln\left(\sqrt{\frac{\sum_{i=0}^{\tau}(r_{t-i}-\bar{r})^2}{\tau}}\right)$ → Calc stock volatility

Mapping

Model Structure

Multimodal Deep Regression Model(MDRM)



LSTM: key information; gradient vanishing; long term

x_j: j_{th} input feature, j_{th} sentence or audio clip

$$f_j = \sigma_g(W_f x_j + U_f h_{j-1} + b_f)$$

$$i_j = \sigma_g(W_i x_j + U_i h_{j-1} + b_i)$$

$$o_j = \sigma_g(W_o x_j + U_o h_{j-1} + b_o)$$

$$c_j = f_j \circ c_{j-1} + i_j \circ \sigma_c(W_c x_j + U_c h_{j-1} + b_c)$$

$$h_j = o_j \circ \sigma_h(c_j)$$

$$Z_j = ReLU(W_z h_j + b_z)$$

Learn inter-modalities dependencies

Improves predicting power

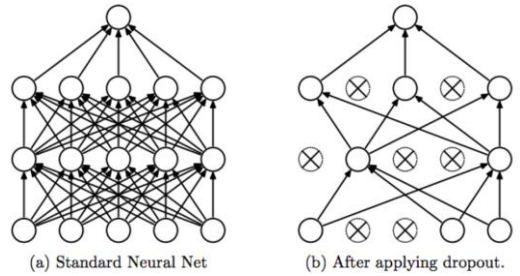
Feature fusion: mask, matrix stack

Experiment

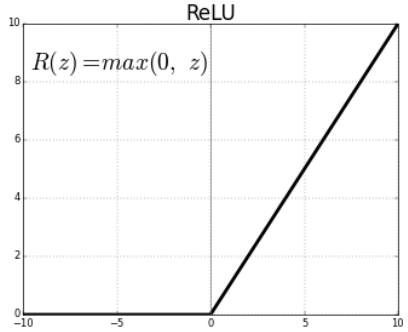
Baselines & Training

Whole Model	Multimodal Feature Fusion	Text Feature Extraction	Modalities
Past Volatility	Simple fusion	TF-IDF bag-of-words + SVR / RBF	Text-only
	bc-LSTM	Word Embeddings + SVR / RBF	Audio-only

Baselines



Dropout



ReLU

Experiment

Result & Discussion

		τ=3	τ=7	τ=15	τ=30
v^{past}		2.986	0.826	0.420	0.231
tf-idf bag-of-words		1.695	0.498	0.342	0.249
word embeddings		1.667	0.549	0.345	0.275
simple fusion		1.722	0.501	0.307	0.233
bc-LSTM (text+audio) (Poria et al., 2017)		1.418	0.436	0.304	0.219
Multimodal Deep Regression Model (MDRM)	text only	1.431	0.439	0.309	0.219
	audio only	1.412	0.440	0.315	0.224
	text+audio	1.371***	0.420***	0.300**	0.217

Table 1: MSE of different models on stock volatility prediction τ-days following the conference call. The * denotes statistical significance compared to MDRM (text only) results under a one-tailed t-test (*** for $p \leq 0.001$ and ** for $p \leq 0.01$)

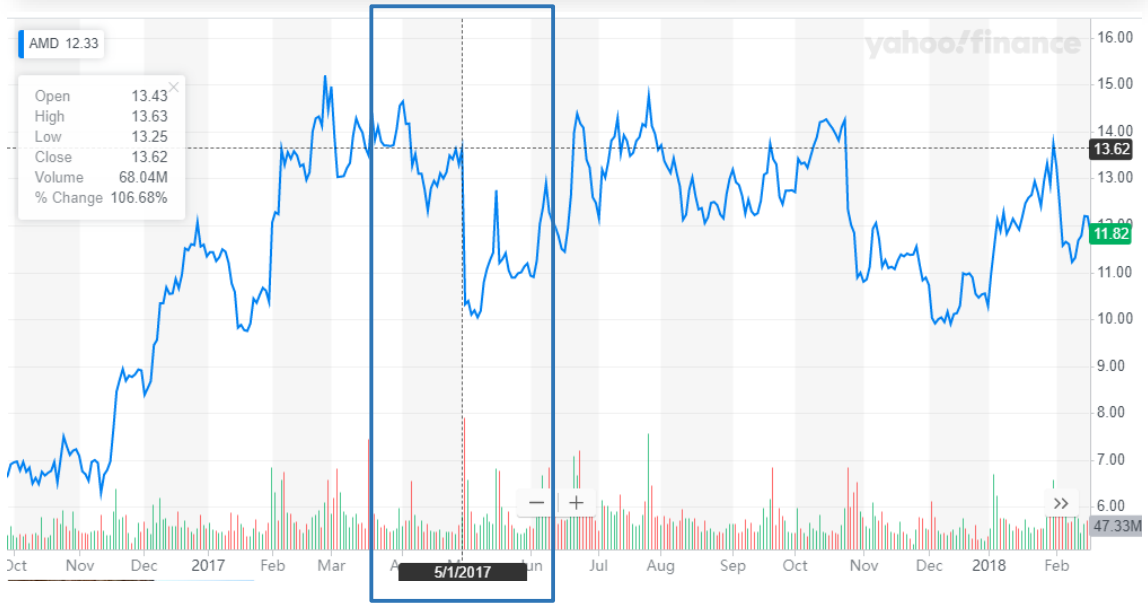
- Model is effective
- Both modalities are helpful
- Some individual vocal cues are important
- Short-term volatility prediction is hard due to PEAD(post earnings announcement drift)
- Margin gain is diminishing in long-term

Case Study: AMD Conference Call Q1 2017

Positive Word is not as Credible as it Sounds

Case 1: *"Overall, from a performance standpoint, the product and the customer engagements are going as we would expect."*

Case 2: *"We have more memory bandwidth."*



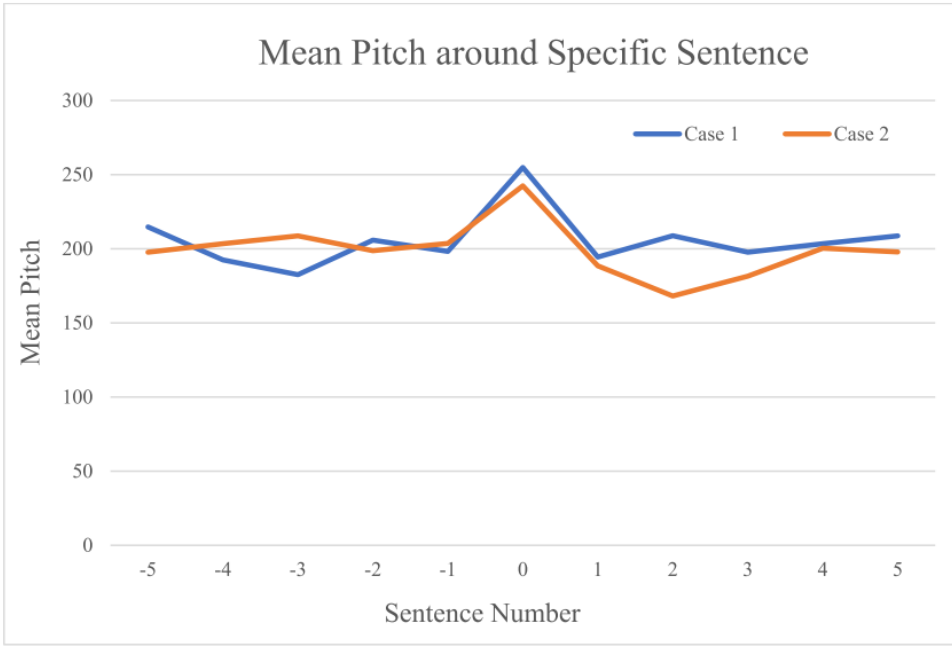


Figure 2: The change of Mean Pitch around specific sentence. Sentence with number 0 is the corresponding Case1 and Case2 sentence described in the paper.

Insights

Explainable Model

Integrate CNN

Baselines

Long/Short-term

Q&A, References

This paper:
https://www.aclweb.org/anthology/P19-1038/

Companies IR & Earnings conference call:
https://www.amd.com/en/press-releases/press-release-2017may01
https://edge.media-server.com/mmc/p/adf6jk9m
https://edge.media-server.com/mmc/p/hcz3or8m

Stock price:
https://finance.yahoo.com/quote/AMD/chart?p=AMD
https://finance.yahoo.com/quote/GS/chart?p=GS

Cai Mingshi / i@unoiou.com / 2019/11/12