

Reading China:

Predicting Policy Change with Machine Learning

Julian TszKin Chan
(Bates White)

Weifeng Zhong
(AEI)

October 25, 2018
GWU

The views expressed here are solely those of our own and do not represent the views of the American Enterprise Institute, Bates White Economic Consulting, or their other employees.

Policy Change Index (PCI) of China

China's industrialization has been a product of gov't direction.

PCI: the first *quantitative* indicator of China's policy priorities.

- A leading indicator, 1951 Q1 – 2018 Q3.
- Can be updated in the future.

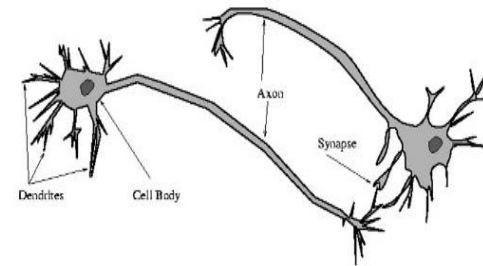
How to predict policy change?

Build a machine learning algorithm to

- “read” the *People's Daily*;
- detect changes in how it prioritizes issues.



Official newspaper, 1946-2018



(Artificial) neural networks

Source of predictive power

People's Daily:
nerve center of China's propaganda system

+

Propaganda often precedes policies.

↓

Detect changes in
newspaper's priorities

≈

Predict changes in
gov't policies

Source of predictive power

People's Daily:
nerve center of China's propaganda system

+

Propaganda often precedes policies.

↓

Detect changes in
newspaper's priorities

≈

Predict changes in
gov't policies

Front page?

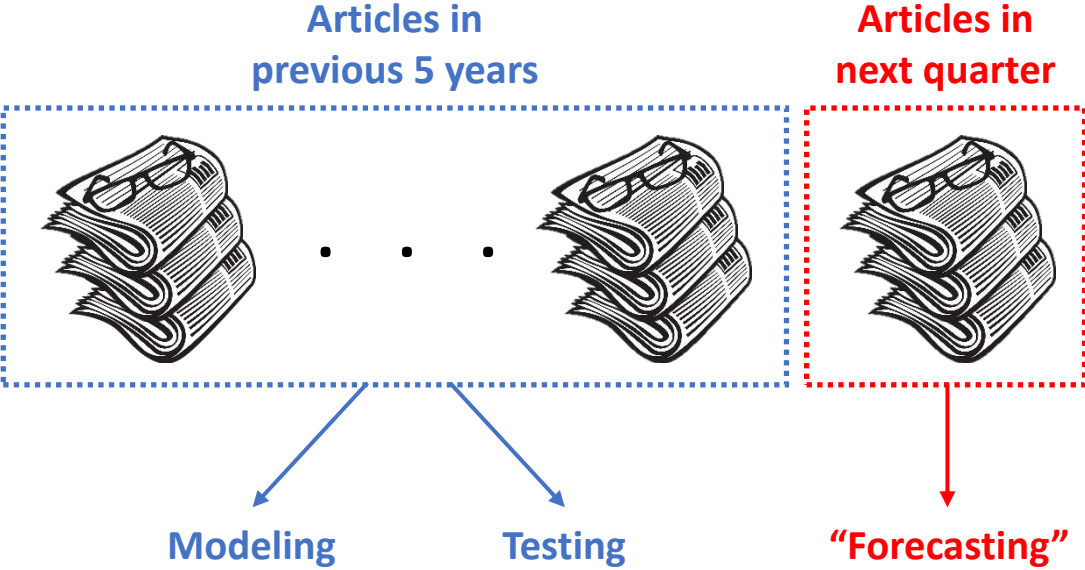


Methodology

Imagine an avid reader of the *People's Daily* who

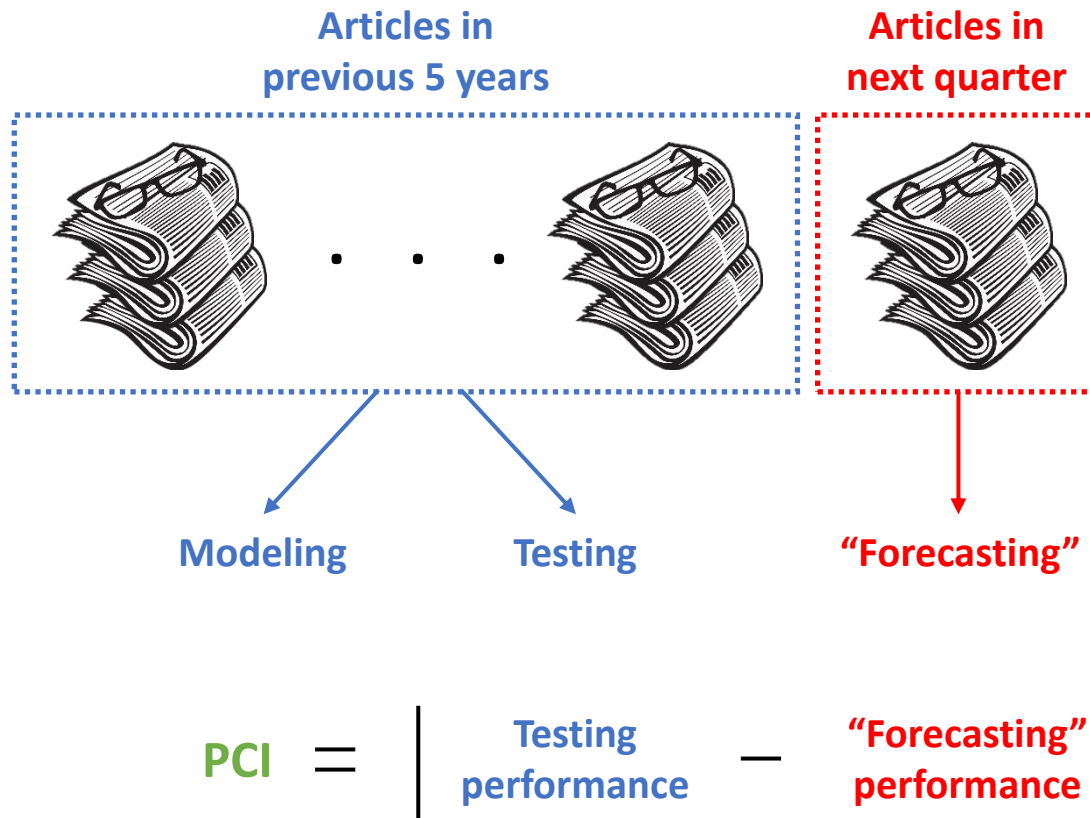
1. reads recent articles;
2. forms a paradigm about front-page content;
3. tests the paradigm on new articles.

Methodology



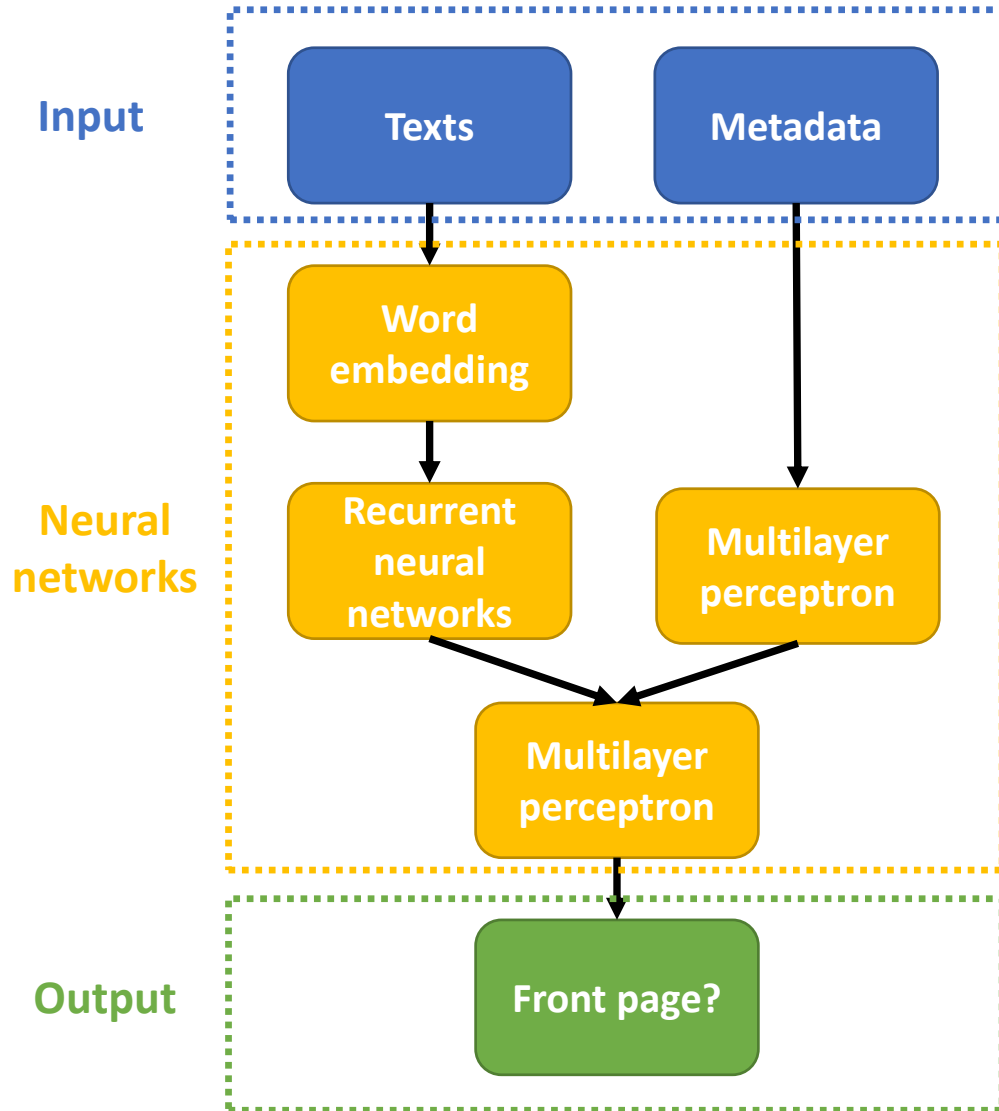
$$PCI = \left| \text{Testing performance} - \text{"Forecasting" performance} \right|$$

Methodology



“Context-free” — it does **not** require an understanding of the Chinese context.

Methodology: modelling



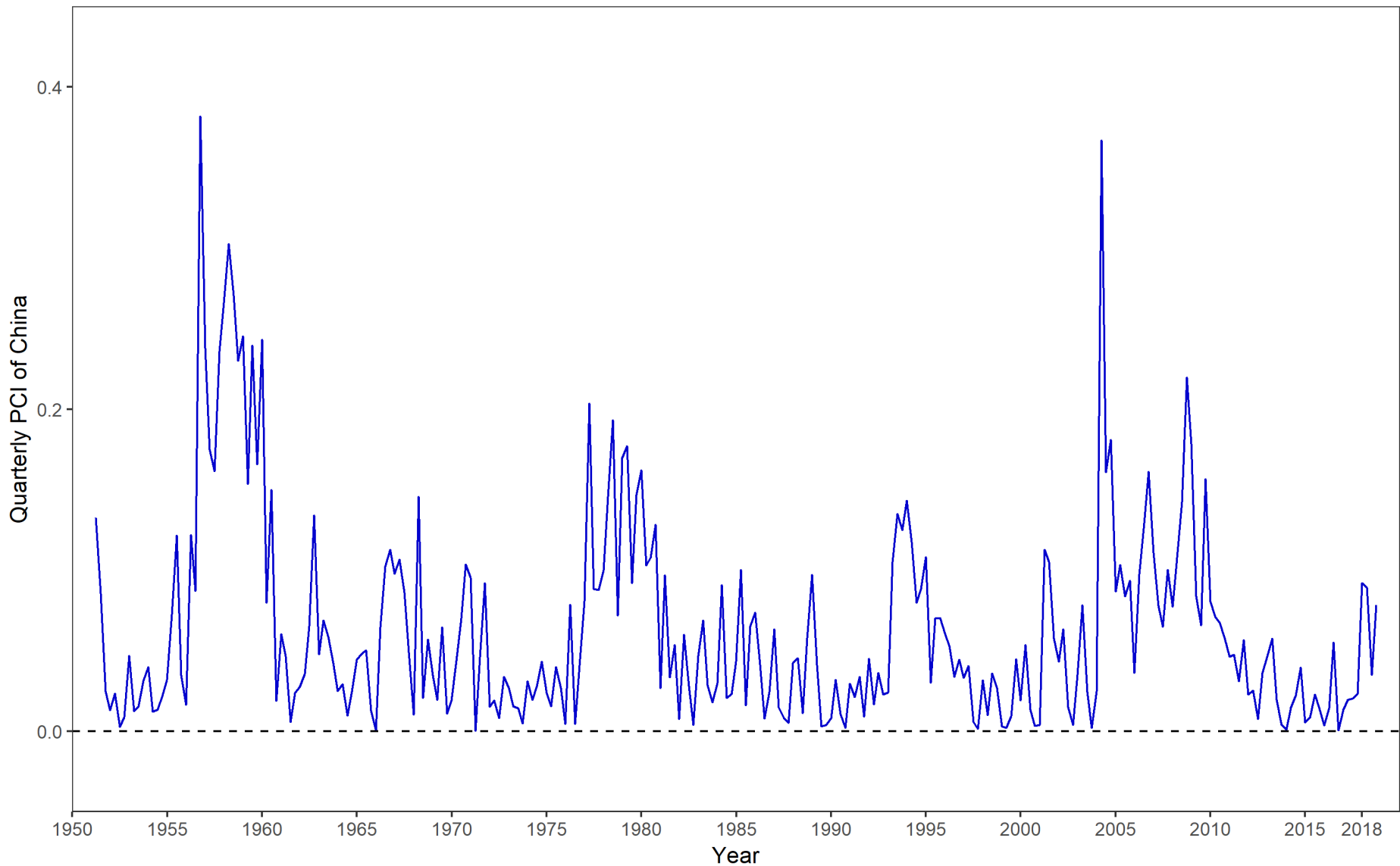
x : each article as an observation.

f : a complicated function.

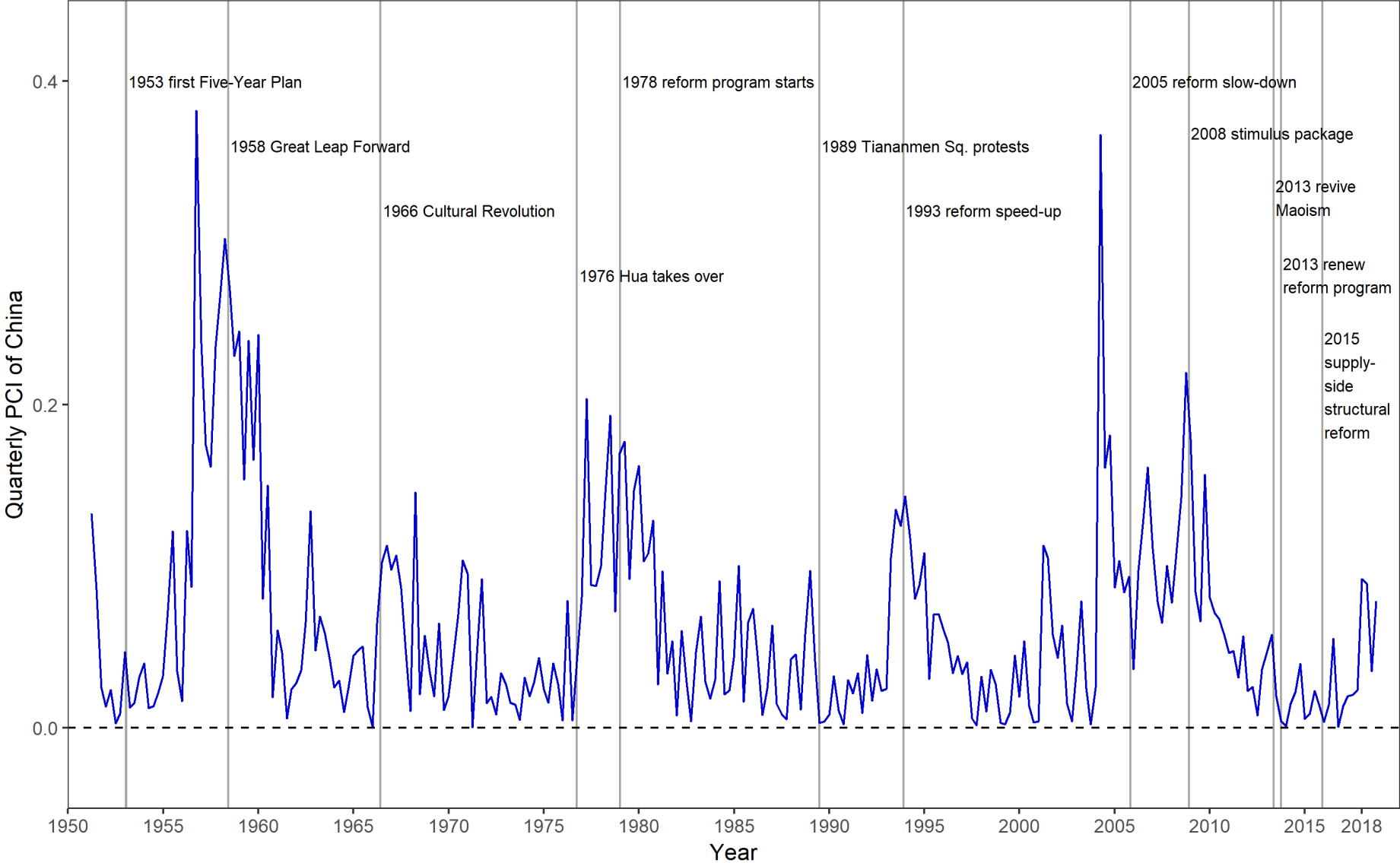
$$y = f(x)$$

Results

Result: PCI



Result: PCI — with ground truth



Understanding substance of change

		Classified on front page?	
		No	Yes
Front page?	No	√	false positives
	Yes	false negatives	√

- Content of *mis*-classified articles has policy substance.

Supervised learning: a digression

Supervised learning

$$\text{mapping} : A \rightarrow B$$

- Task: to learn the mapping from $\{a_i, b_i\}_{i \in \text{training}}$.
- Goal: from $\{a_j\}_{j \in \text{new}}$, to predict $\{b_j\}_{j \in \text{new}}$.
- Challenge: need a lot training data.

The newspaper problem: an *infeasible* approach

$$g : \{\textit{articles}\} \times \{\textit{front page?}\} \rightarrow \{\textit{priorities}\}$$

- With the learned function g :
 - $g(\text{"pvt sector is important"}, \text{front page}) = \text{pro-reform};$
 - $g(\text{"should protect SOEs"}, \text{not front page}) = \text{pro-reform};$
 - $g(\text{"central planning is awesome"}, \text{front page}) = \text{anti-reform}; \dots$

The newspaper problem: an *infeasible* approach

$$g : \{\textit{articles}\} \times \{\textit{front page?}\} \rightarrow \{\textit{priorities}\}$$

- With the learned function g :
 - $g(\text{"pvt sector is important"}, \text{front page}) = \text{pro-reform};$
 - $g(\text{"should protect SOEs"}, \text{not front page}) = \text{pro-reform};$
 - $g(\text{"central planning is awesome"}, \text{front page}) = \text{anti-reform}; \dots$
- But where are the training data?

The newspaper problem:

a feasible approach

- Consider an “inverse function” f of g :

$$f : \{\textit{articles}\} \times \{\textit{priorities}\} \rightarrow \{\textit{front page?}\}$$

The newspaper problem: a feasible approach

- Consider an “inverse function” f of g :

$$f : \{\textit{articles}\} \times \{\textit{priorities}\} \rightarrow \{\textit{front page?}\}$$

- Condition f on a given priority s :

$$f_s : \{\textit{articles}\} \times \{\textit{priority } s\} \rightarrow \{\textit{front page?}\}$$

The newspaper problem: a feasible approach

- Consider an “inverse function” f of g :

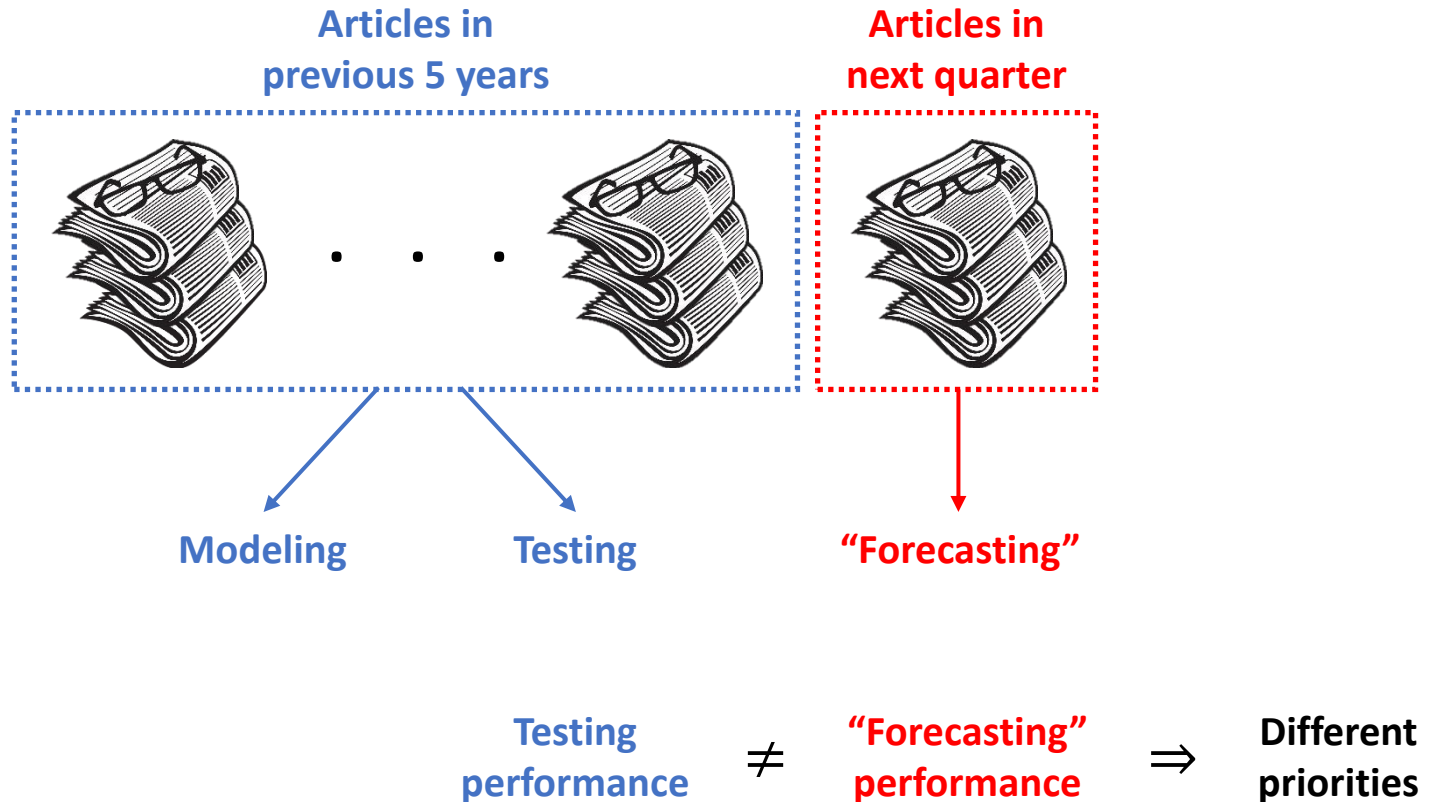
$$f : \{\text{articles}\} \times \{\text{priorities}\} \rightarrow \{\text{front page?}\}$$

- Condition f on a given priority s :

$$f_s : \{\text{articles}\} \times \{\text{priority } s\} \rightarrow \{\text{front page?}\}$$

- Lots of training data to learn each f_s .
- $f_s \neq f_t \Rightarrow$ priority $s \neq$ priority t .
- “Context-free!”

The newspaper problem: a feasible approach



Potential applications

App 1: PCI of other countries

Predicting other (ex-)Communist regimes' policies:

- USSR's *Pravda*
- GDR (East Germany)'s *Neues Deutschland*
- DPRK (North Korea)'s *Rodong Sinmun*
- Cuba's *Granma*
- Vietnam's *Nhân Dân*

App 2: measuring centralization

Measuring degree of centralization in China:

- Replicate the same analysis on regional official newspapers.
- Compare $PCI_{Central}$, $PCI_{Region 1}$, $PCI_{Region 2}$, ...
- Conforming to $PCI_{Central}$ \rightsquigarrow Toeing the party line

App 3: measuring media bias

Measuring media bias in the US:

- Replicate the same analysis on US newspapers.
- Compare " PCI "_{WaPo}, " PCI "_{NYT}, " PCI "_{WSJ}, ...
- Divergence among " PCI "s → Polarization in media

App 4: predicting vote change

Predicting vote change in legislation:

- *People's Daily* articles → **Legislators' public statements**
- Page numbers of articles → **Legislators' names**
- What if Sen. *A*'s statement is mistaken as Sen. *B*'s?

Source of predictive power:

- Political necessity to justify vote changes by making different statements *in advance*.

App 5: predicting ideology change

Predicting changes in judges' ideological leaning:

- *People's Daily* articles → **Judges' judicial opinions**
- Page numbers of articles → **Judges' names**
- What if Judge *C*'s opinion is mistaken as Judge *D*'s?

Source of predictive power:

- Changes in ideological leaning tend to be gradual yet long-lasting.

Takeaways

- The first index to predict China's "next big things."
- A new way to uncover hidden patterns from trivial labels.
- A wide range of potential applications.

Interested in DIY?

- Website: www.policychangeindex.com
- Research paper: www.policychangeindex.com/Reading_China.pdf
- Source code: www.github.com/open-source-economics/PCI

Contact us:

- Julian TszKin Chan: ctszkin@gmail.com
- Weifeng Zhong: weifeng@weifengzhong.com

Questions?