



keeping informed

->



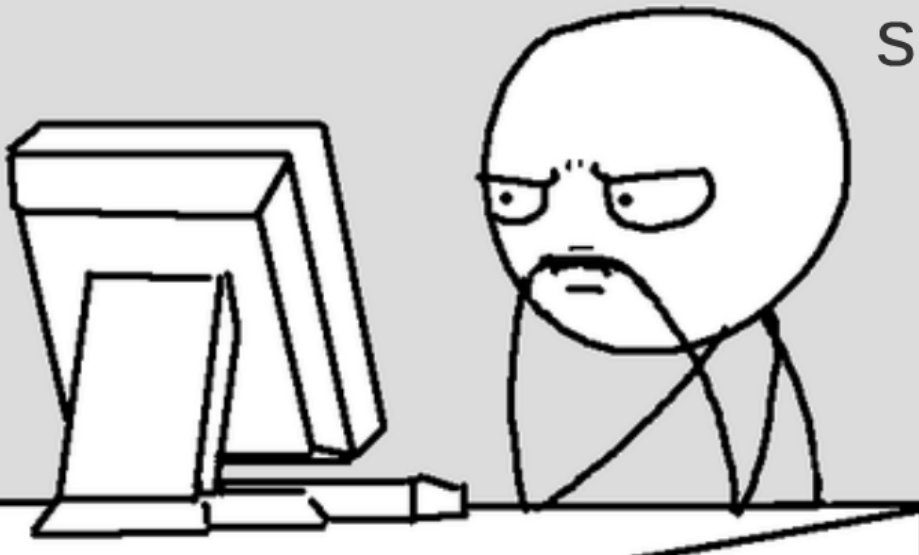
life is
depressing

war!

disease!

surveillance!

famine!



SHOP
MARCJACOBS.COM
LITTLE MARC

The New York Times

Sunday, April 27, 2014 | Today's Paper | Personalize Your Weather |



WORLD U.S. NEW YORK BUSINESS OPINION SPORTS SCIENCE ARTS FASHION & STYLE VIDEO

All Sections

U.S. Weighs Harder Line on Russia Than European Allies

By PETER BAKER and C. J. CHIVERS
8:43 PM ET

The White House has become entangled in a debate over how much emphasis to put on unity with European allies reluctant to take stronger economic actions against Moscow, while pro-Russian forces in Ukraine continue to defy demands.

Pro-Russian Forces Free One Observer in Eastern Ukraine

By C. J. CHIVERS and NOAH SNEIDER

The self-appointed mayor of Slovyansk presented eight members of a military observer mission at a news conference on Sunday, calling them "prisoners of the situation."

Two Popes Become Saints in Historic Canonizations



Ko Sasaki for The New York Times

Fukushima Evacuees Are Hesitant to Return

By MARTIN FACKLER 6:44 PM ET

Three years after the nuclear accident, the Japanese government has reopened a town that many do not want to return to. Above, bags of radiation-contaminated waste.

Clippers Take to the Court, but in Protest

By JOHN BRANCH 41 minutes ago

The players made a statement a day after racially charged remarks attributed to Donald Sterling, the Clippers' owner, by wearing their shirts inside out before playing the Warriors.

98 Comments

On Pro Basketball: In Complex Mess, Who Has Leverage

Mayor Who Likes to Share His View From the Top

By ADAM NAGOURNEY 24 minutes ago

Personal photographs taken by Mayor Eric Garcetti shed light on Los Angeles and offer a glimpse into the subdued and slightly offbeat style

NEWS ANALYSIS

A Birdbath-Size Hole and the Future of Golf

By CHARLES McGRATH

The sport's stewards might want to think twice before tinkering with dimensions.



OPINION

The Global Diabetes Epidemic

By KASIA LIPSKA

In the West, it's a manageable chronic condition. In places like India, it's a different story.



- Editorial: The Koch Attack on Solar Energy
- Bruni: Larry Kramer
- Dowd: Slaves to Prejudice
- Douthat: The Pope's Call
- Friedman: Who Will Influence Whom?
- Kristof: Religion for \$1,000, Alex

Today's Times Insider

Behind the scenes of The New York Times

- Gretchen Morgenson on Wall Street
- Introducing Times Insider



2x Dining & Entertainment

hypothesis: the words in a document
predict the document's "joy score"*

*what is joy?

*Item: Record Low Set In Deaths From Traffic, Mayor
Says*

Item: Pierce Is Critically Injured On Halfpipe

Item: Rush Limbaugh Leaves Hospital

data selection:

source: New York Times archives
timeframe: Jan 1, 2010 - Jan 1, 2012*

data preparation:

1000 initial corpus
833 -167 de-dupe headlines
827 -6 drop null values
748 -79 after removing short headlines (<20 chars)

```
Out[43]: 0      Pearce Is Critically Injured On Halfpipe  
         1              What's On Today  
         2 Celebrations in Times Square Despite Troubled ...  
         3      A Variety of Stories On Many Screens  
         5      Paid Notice: Deaths  DEIN, GITTIE
```

data scoring:

- by hand :-/
- optimize for individual

```
articles[['snippet', 'lead_paragraph', 'abstract', 'headline', 'pub_date', 'word_count']]
```

	snippet	lead_paragraph	abstract	headline	pub_date	word_count
0	The Royal Mail steamship Canada, Captain HARRI...	The Royal Mail steamship Canada, Captain HARRI...	Progress of the Rebellion in	{"main"=>"THREE DAYS LATER FROM EUROPE; ARRIVA...	1851-10-01T00:03:58Z	2152
1	NaN	NaN	NaN	{"main"=>"MARINE INTELLIGENCE.; Cleared. Arriv...	1851-10-01T00:03:58Z	1281
2	MY DEAR MR. GRINNELL: I write to you in much a...	MY DEAR MR. GRINNELL: I write to you in much a...	Lady Franklins letter	{"main"=>"Letter from Lady Franklin to Mr. Gri...	1851-10-02T00:03:58Z	990
3	The Journal of Commerce published the followin...	The Journal of Commerce published the followin...	Letter of Andw. Stephenson on the European view	{"main"=>"The Cuban Question in Europe--Letter...	1851-10-02T00:03:58Z	991
4	Any one who comes to Munich with the purpose o...	Any one who comes to Munich with the purpose o...	NaN	{"main"=>"SKETCHES OF EUROPE.; Munich--its Gal...	1851-10-02T00:03:58Z	2165
5	TRIPLER HALL.--Fifth concert by Catharine Haye...	TRIPLER HALL.--Fifth concert by Catharine Haye...	NaN	{"main"=>"NEW-YORK CITY.; AMUSEMENTS THIS EVEN...	1851-10-02T00:03:58Z	2426
6	The Queen returned from Ischl, yesterday and p...	The Queen returned from Ischl, yesterday and p...	NaN	{"main"=>"PRUSSIA.", "kicker"=>"1"}	1851-10-03T00:03:58Z	444
7	NaN	NaN	NaN	{"main"=>"MARINE INTELLIGENCE.", "kicker"=>"1"}	1851-10-03T00:03:58Z	1260
8	NaN	NaN	NaN	{"main"=>"ONE DAY LATER FROM EUROPE.", "kicker...	1851-10-03T00:03:58Z	133
9	NaN	NaN	NaN	{"main"=>"THREE DAYS LATER FROM EUROPE; ARRIVA...	1851-10-03T00:03:58Z	30

web_url
snippet
lead_paragraph

abstract
print_page
blog

source
multimedia

headline

keywords

pub_date

document_type

news_desk

section_name

subsection_name

byline

type_of_material

_id

word_count

images
video
author

Finding patterns, exploring...

1) TFIDF transform

- fast
- scalable
- great reputation for text problems
- joyful:1, everything else:0

2) Random Forest Classifier

- avoid overfitting
- speed not an issue
- no distribution assumptions
- find best term features?

3) Naive Bayes

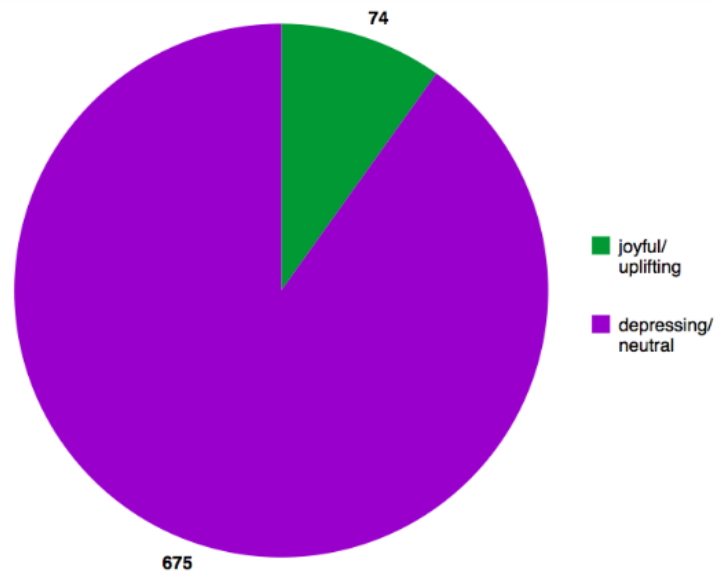
- fast
- scalable (linear with corpus)
- iterate & search for best feature set:
 - headline text
 - first paragraph text
 - headline & first paragraph text

[0.908 0.8965 0.891]

error: $\sim .90$



false positive rate: $\sim 100\%$



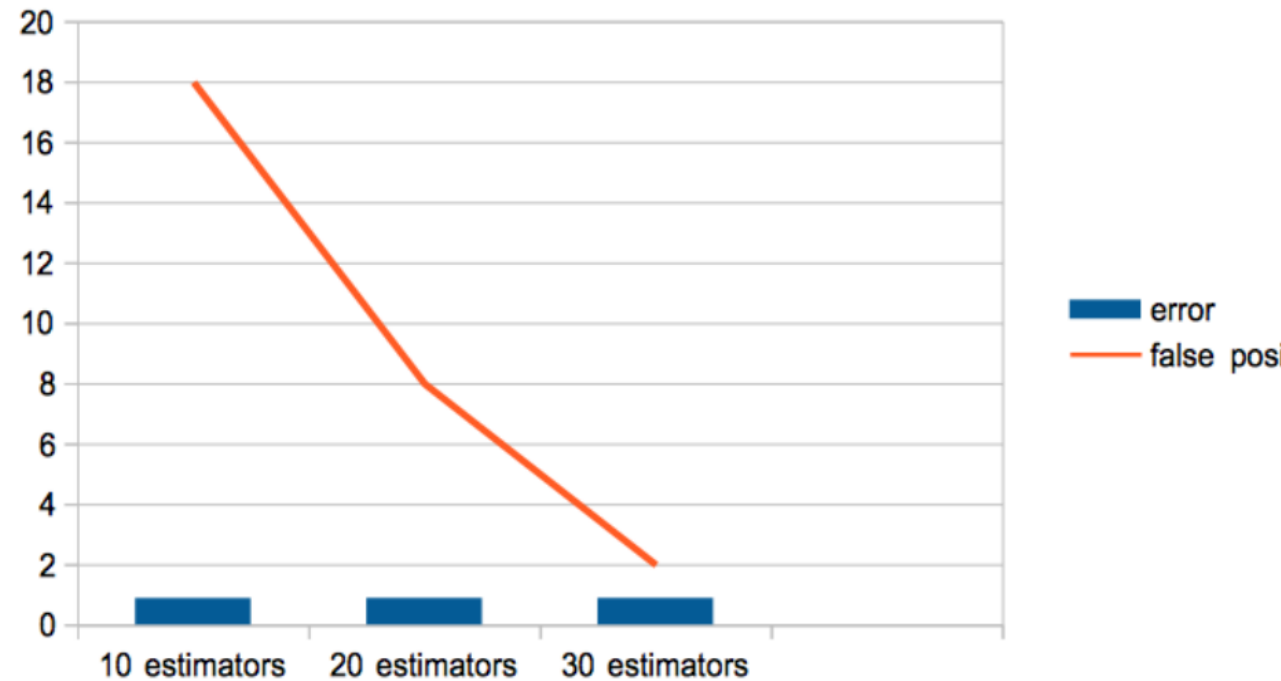
-> more data ->

results

Random Forest

error: .879

false positive rate:
 $2/94 = .021$



Naive Bayes

error: .870

false positive rate: $4/94 = 4.6\%$

SHOP
MARC JACOBS.COM
LITTLE MARC

The New York Times

Sunday, April 27, 2014 Today's Paper Personalize Your Weather



WORLD U.S. NEW YORK BUSINESS OPINION SPORTS SCIENCE ARTS FASHION & STYLE VIDEO

All Sections

U.S. Weighs Harder Line on Russia Than European Allies

By PETER BAKER and C. J. CHIVERS
8:43 PM ET

The White House has become entangled in a debate over how much emphasis to put on unity with European allies reluctant to take stronger economic actions against Moscow, while pro-Russian forces in Ukraine continue to defy demands.

Pro-Russian Forces Free One Observer in Eastern Ukraine

By C. J. CHIVERS and NOAH SNEIDER

The self-appointed mayor of Slovyansk presented eight members of a military observer mission at a news conference on Sunday, calling them "prisoners of the situation."

Two Popes Become Saints in Historic Canonizations



Ko Sasaki for The New York Times

Fukushima Evacuees Are Hesitant to Return

By MARTIN FACKLER 6:44 PM ET

Three years after the nuclear accident, the Japanese government has reopened a town that many do not want to return to. Above, bags of radiation-contaminated waste.

Clippers Take to the Court, but in Protest

JOHN BRANCH 41 minutes ago

The players made a statement a day after racially charged remarks attributed to Donald Sterling, the Clippers' owner, by wearing their shirts inside out before playing the Warriors.

98 Comments

On Pro Basketball: In Complex Mess, Who Has Leverage

Mayor Who Likes to Share His View From the Top

By ADAM NAGOURNEY 24 minutes ago

Personal photographs taken by Mayor Eric Garcetti shed light on Los Angeles and offer a glimpse into the subdued and slightly offbeat style



The Opinion Pages

NEWS ANALYSIS

A Birdbath-Size Hole and the Future of Golf

By CHARLES McGRATH

The sport's stewards might want to think twice before tinkering with dimensions.



OPINION

The Global Diabetes Epidemic

By KASIA LIPSKA

In the West, it's a manageable chronic condition. In places like India, it's a different story.



- Editorial: The Koch Attack on Solar Energy
- Bruni: Larry Kramer
- Dowd: Slaves to Prejudice
- Douthat: The Pope's Call
- Friedman: Who Will Influence Whom?
- Kristof: Religion for \$1,000, Alex

Today's Times Insider

Behind the scenes of The New York Times

- Gretchen Morgenson on Wall Street
- Introducing Times Insider



2x Dining & Entertainment

SHOP
MARCJACOBS.COM
LITTLE MARC

The New York Times

Sunday, April 27, 2014 Today's Paper Personalize Your Weather



WORLD U.S. NEW YORK BUSINESS OPINION SPORTS SCIENCE ARTS FASHION & STYLE CROSSWORDS VIDEO TIMES PREMIER All Sections

Celebrations in Times Square Despite Troubled Times

By PETER BAKER and C. J. CHIVERS
8:43 PM ET

The White House has become entangled in a debate over how much emphasis to put on unity with European allies reluctant to take stronger economic actions against Moscow, while pro-Russian forces in Ukraine continue to defy demands.

Record Low Set in Deaths From Traffic, Mayor Says

By C. J. CHIVERS and NOAH SNEIDER

New York City has re-engineered streets with pedestrian plazas, bicycle lanes and time-equipped crosswalk signals in recent years.

Two Popes Become Saints in Historic Canonizations



Ko Sasaki for The New York Times

Tax Benefits From Options as Windfall for Businesses

By MARTIN FACKLER 6:44 PM ET

Companies that granted stock options after the 2008 stock market collapse are benefiting from those awards too, in the form of tax savings.

In Pennsylvania, a Quick Shot of Peace, on a Budget

By JOHN BRANCH 40 minutes ago

A Jesuit center provides an Arcadian setting for a quick shot of affordable quiet.

93 Comments

On Pro Basketball: In Complex Mess, Who Has Leverage



A Young Mayor Who Likes to Share His View From the Top

By ADAM NAGOURNEY 23 minutes ago

Personal photographs taken by Mayor Eric Garcetti shed light on Los Angeles and new ways for politicians to use social media.



The Opinion Pages

NEWS ANALYSIS

A Birdbath-Size Hole and the Future of Golf

By CHARLES McGRATH

The sport's stewards might want to think twice before tinkering with dimensions.



OPINION

The iKnight: Apple's Jonathan Ive Gets British Honor

By KASIA LIPSKA

That will be Sir Jony, from now on.



- Editorial: The Koch Attack on Solar Energy
- Bruni: Larry Kramer
- Dowd: Slaves to Prejudice
- Douthat: The Pope's Call
- Friedman: Who Will Influence Whom?
- Kristof: Religion for \$1,000, Alex

Today's Times Insider

Behind the scenes of The New York Times

- Gretchen Morgenson on Wall Street
- Introducing Times Insider



2x Dining & Entertainment

thankyou
PREFERRED

Lessons learned

- know what you want to optimize
- the most accurate model may not be the best IRL
- "joy", "uplifting" hard to define
- more intelligent weighting
- more, less skewed data

Planned Improvements

- stemming
- more data sources
- hybrid individual - community model

husband
american
fact
help
ohio
newspapers
season
try
throwing
park

second time:
businesses
14
impact
today
puzzle
journalist
mass
talking
celebration
candidate