

Text Classification for RTB

Defining Targeted in Targeted Digital Advertising

Text Classification for RTB



LATEST NATIONAL ▾ WORLD VIDEO PHOTOS WEATHER MORE ▾

Orchids of the Hastings

prime7n1 on September 12, 2022, 10:00 am



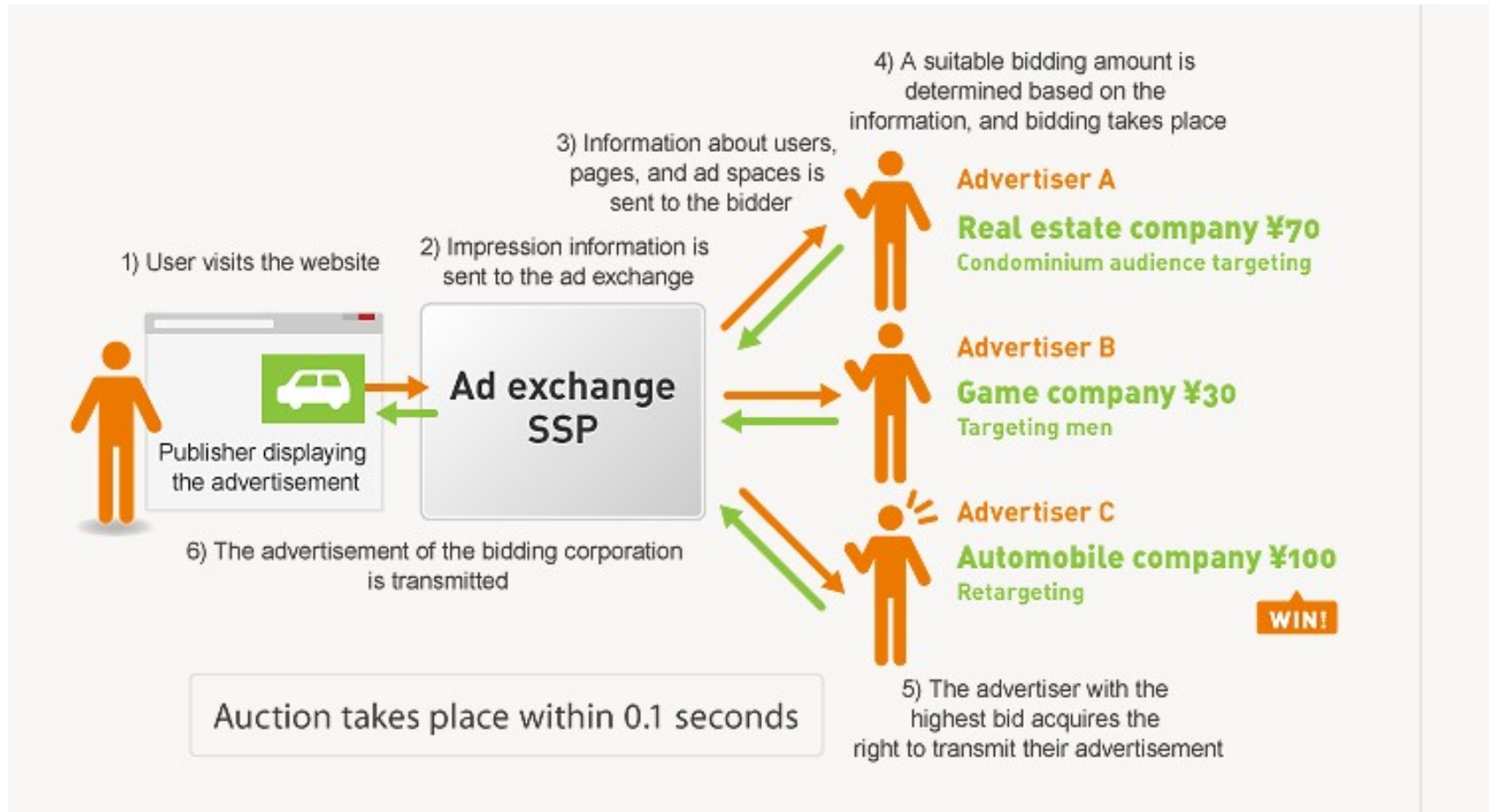
FOLLOW US



WEATHER



Text Classification for RTB



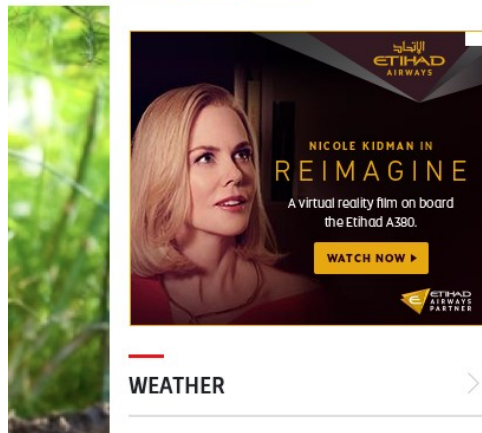
Text Classification for RTB

- Age
- Gender
- General Interests
- Current Purchase Intent



Text Classification for RTB

Buy Third Party Data



Develop Custom Analytics

Text Classification for RTB

The segments we buy:

- Limited Number of Segments
- Hard to sample from a segment
- Those companies do not have info for each and every user
- It adds cost to advertisement
- How do they know? Whose data to purchase?

Text Classification for RTB

Consistency within Dataset

- Is yesterday's male of 55 years old still male and over 55 today?

Comparison of datasets (from different data providers)

- Overlaps: how many new users each provider gives? How many users are in common? Do they belong to same segments?

Text Classification for RTB

Very **!!!EXPENSIVE!!!**



Not only you pay for data, everyone is after the same user at the ad placement auction.

Text Classification for RTB

Advertisement is not soccer!!!



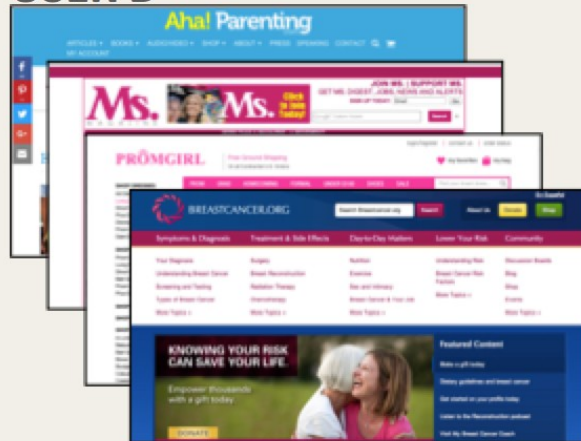
Text Classification: Defining Targeted in Targeted Digital Advertising

USER A



1. <http://www.bgames.com/shooting-games/>
2. <http://www.menshealth.com/fitness/best-lower-abs-exercise>
3. <http://www.bbc.com/news/world-us-canada-33559853>
4. <http://www.paulsmith.co.uk/uk-en/shop/mens/accessories/ties>

USER B



1. <http://www.ahaparenting.com/parenting-tools/intelligent-creative-child>
2. <http://www.msmagazine.com>
3. http://www.promgirl.com/shop/styles_and_trends/long
4. <http://www.breastcancer.org>

Text Classification for RTB

Topics

gene 0.04
dna 0.02
genetic 0.01
...

life 0.02
evolve 0.01
organism 0.01
...

brain 0.04
neuron 0.02
nerve 0.01
...

data 0.02
number 0.02
computer 0.01
...

Documents

Seeking Life's Bare (Genetic) Necessities

COLD SPRING HARBOR, NEW YORK—How many genes does an organism need to survive? Last week at the genome meeting here,* two genome researchers with radically different approaches presented complementary views of the basic genes needed for life. One research team, using computer analyses to compare known genomes, concluded that today's organisms can be sustained with just 250 genes, and that the earliest life forms required a mere 128 genes. The other researcher mapped genes in a simple parasite and estimated that for this organism, 800 genes are plenty to do the job—but that anything short of 100 wouldn't be enough.

Although the numbers don't match precisely, those predictions

are not all that far apart," especially in comparison to the 75,000 genes in the human genome, notes Siv Andersson, a molecular biologist at Uppsala University in Sweden, who arrived at the 800 number. But coming up with a consensus answer may be more than just a genetic numbers game, particularly as more and more genomes are completely mapped and sequenced. "It may be a way of organizing any newly sequenced genome," explains Arcady Mushegian, a computational molecular biologist at the National Center for Biotechnology Information (NCBI) in Bethesda, Maryland. Comparing an

Stripping down. Computer analysis yields an estimate of the minimum modern and ancient genomes.

* Genome Mapping and Sequencing, Cold Spring Harbor, New York, May 8 to 12.

SCIENCE • VOL. 272 • 24 MAY 1996

Topic proportions and assignments

Text Classification for RTB

- The **text classification** itself is a BigData problem – quality depends on the amount of training data
- **User classification** [after discovering what they read] is definitely a BigData problem. The amount of people who browse on-line is huge. So is the number of logs we have to process!!!

Text Classification for RTB

user_id, ts1, url1,

user_id, ts2, url2,

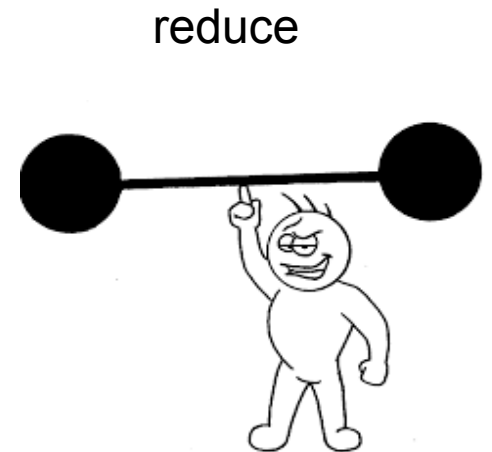
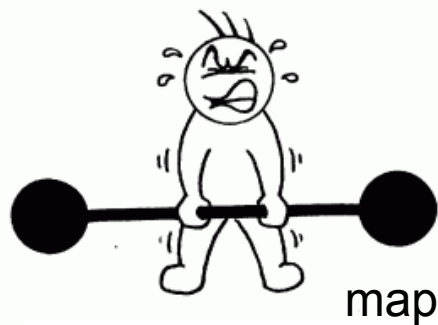
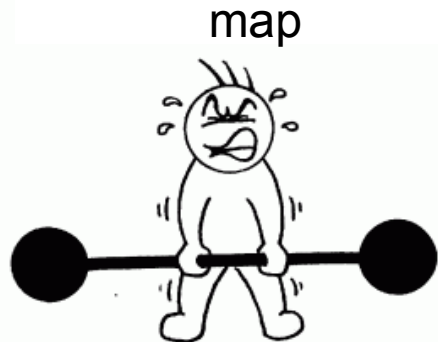
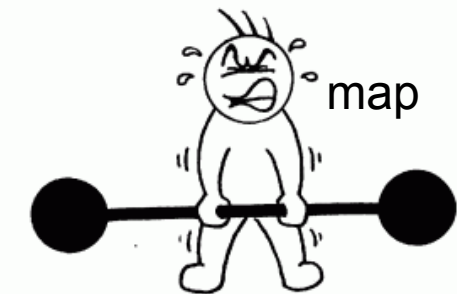
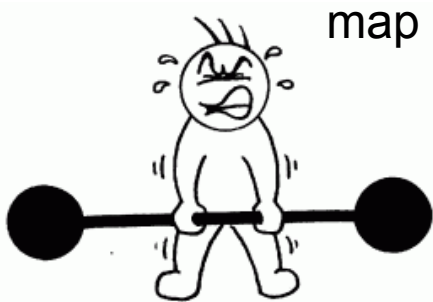
user_id, ts3, url3,

Url → category

=>

user_id, {ts1 category1}, {ts2 category2}, {ts3
category3}, ...

Text Classification for RTB



Text Classification for RTB

- Store lots of data
- Look after a cluster of computers
- Support a number of services: Text Classification Service, User Labelling service, Segment Creation, Reporting

Results

- Ability to clearly define segments
- Ability to meaningfully decrease or increase the size of a segment
- Ability to build custom segments per advertisers's request
- Ability to provide better reporting to customers
- Not spending money on third party data