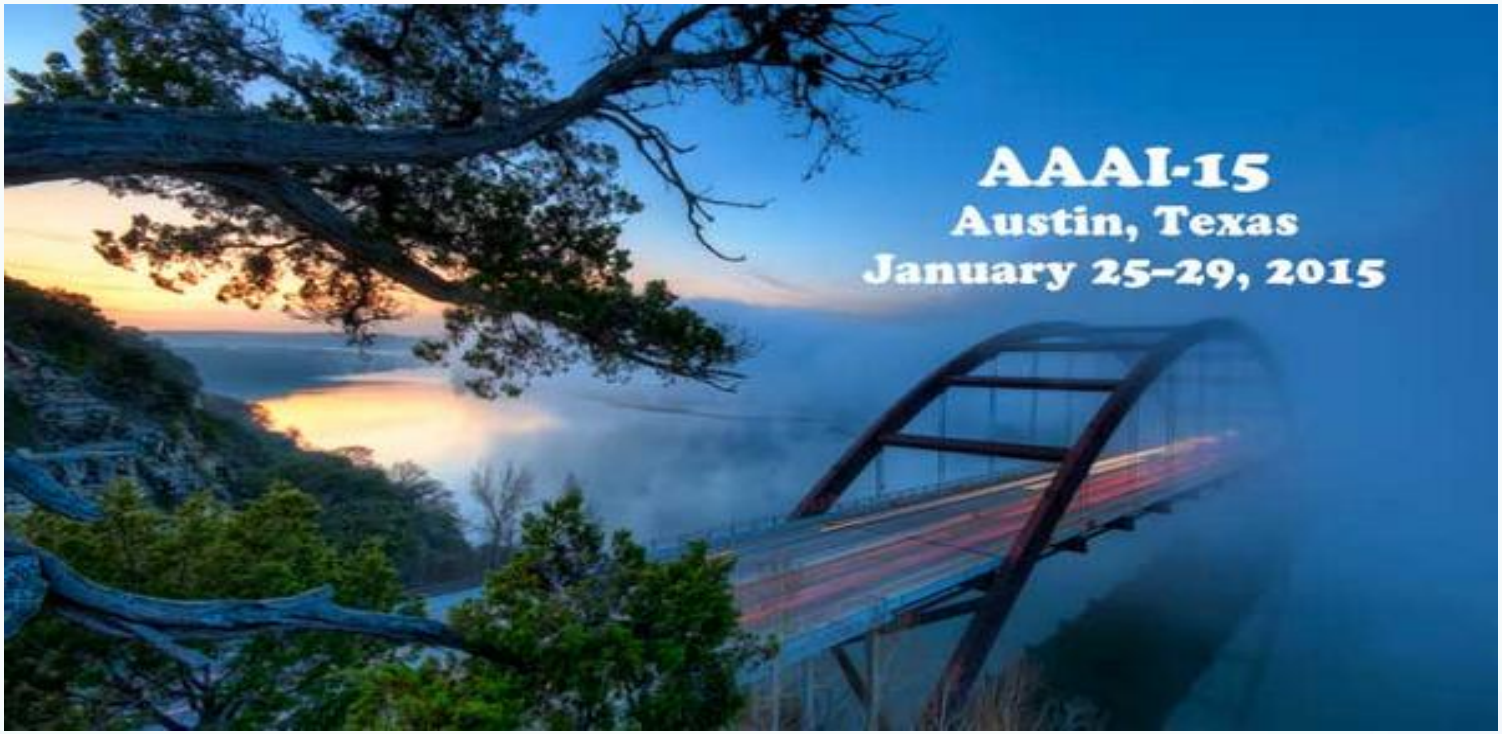
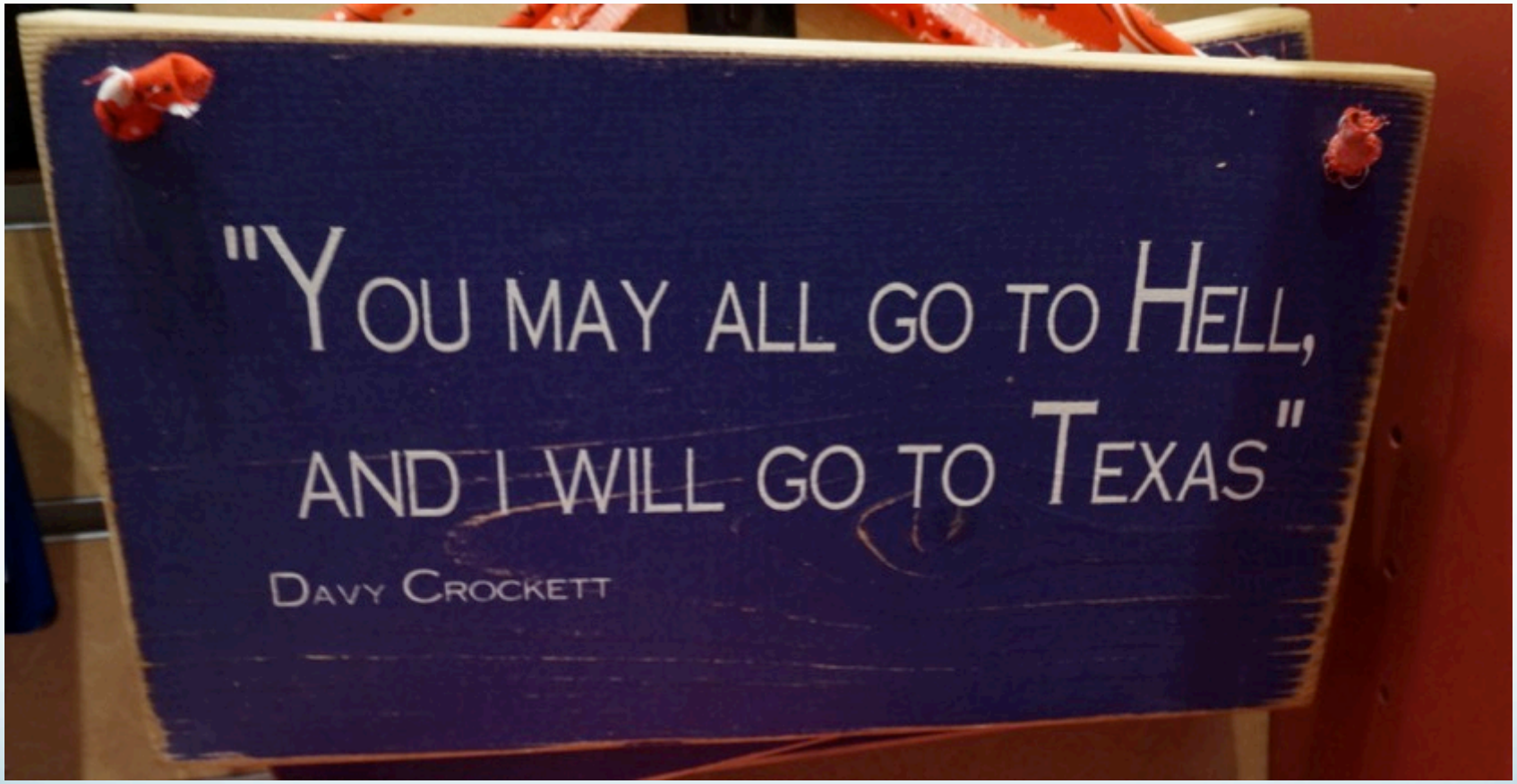


Debrief by Tao Chen
Feb 27, 2015

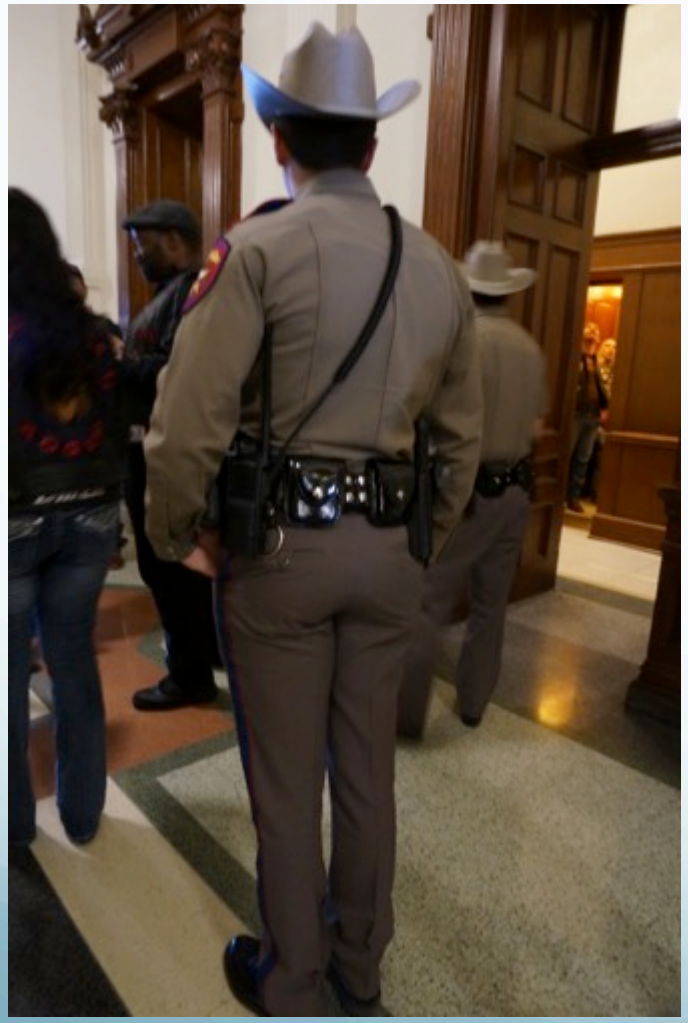
Austin, Texas, USA

Texas: The Lone Star State


"YOU MAY ALL GO TO HELL,
AND I WILL GO TO TEXAS"
DAVY CROCKETT

Before I went

When I was there



Texas State Capitol


GEORGE W. BUSH

Colorado River

University of Texas, Austin

Reception at UT, Austin

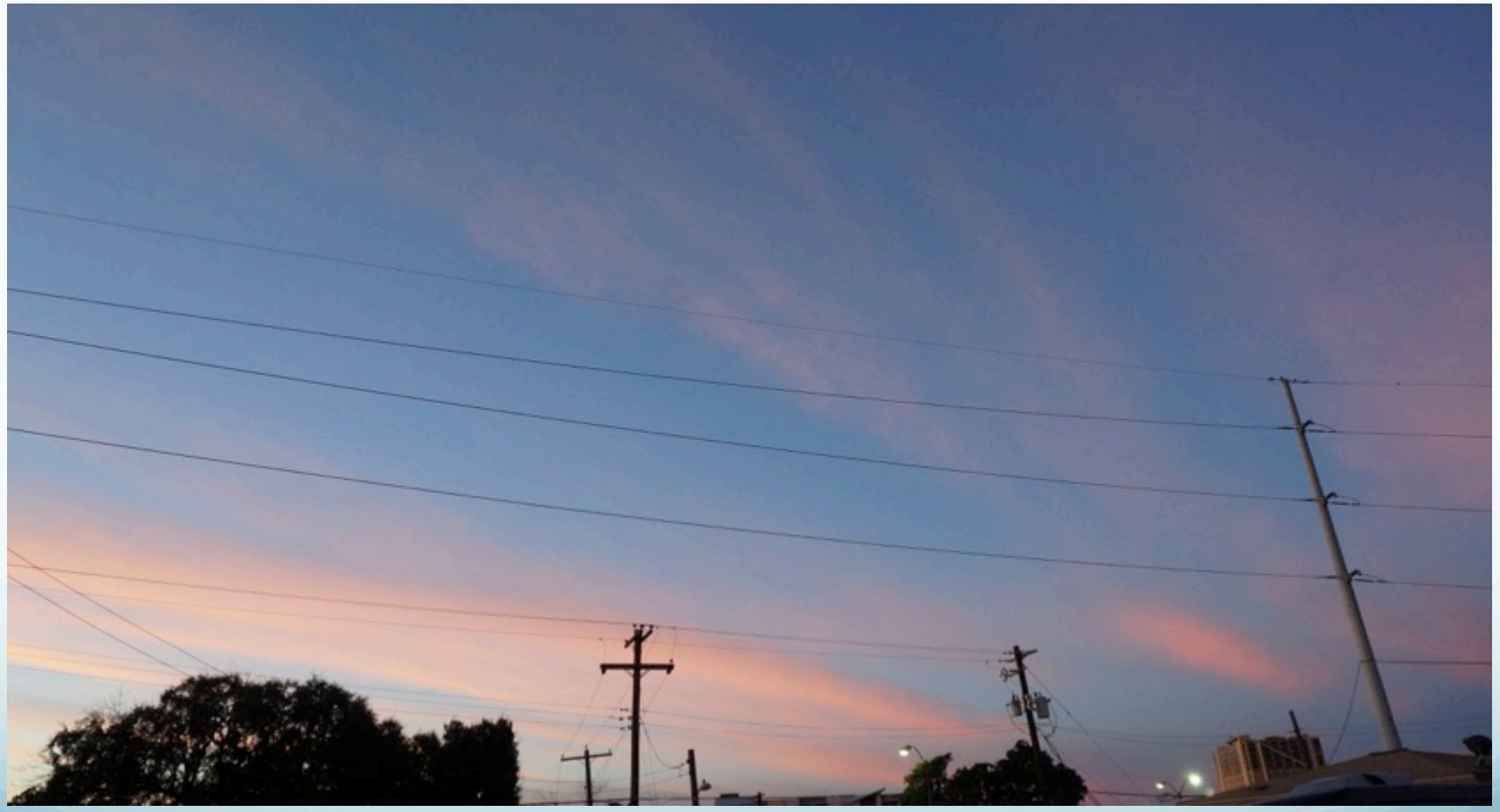
Big Picture of AAAI

- Information about main technical track
  - 1991 submissions (1406 submission in AAAI-14)
  - 539 accepted papers (=27% acceptance rate)
  - AAAI-15 is 5.5 days (one day longer than AAAI-14)
  - First winter AI conference

- Tracks
  - AI and the Web (7 sessions)
  - Natural Language Processing (4 sessions)
  - Machine Learning (9 sessions)
  - Vision (3 sessions)
  - Traditional AI: Cognitive Systems, Computational Sustainability, Game Theory, Multiagent Systems, etc

https://twitter.com/maidylm/status/560542250195619840

Tight Schedule: 8:30am – 8:30pm



Talks Given by Senior Members

- Senior Member Blue Sky Talks
- What's Hot Talks
- Classic Paper Talk
- Panel Discussions

Breakfast with Champions

Lunch with an AAAI Fellow

Murray Campbell, Father of Deep Blue

Robots are everywhere!

Best Papers

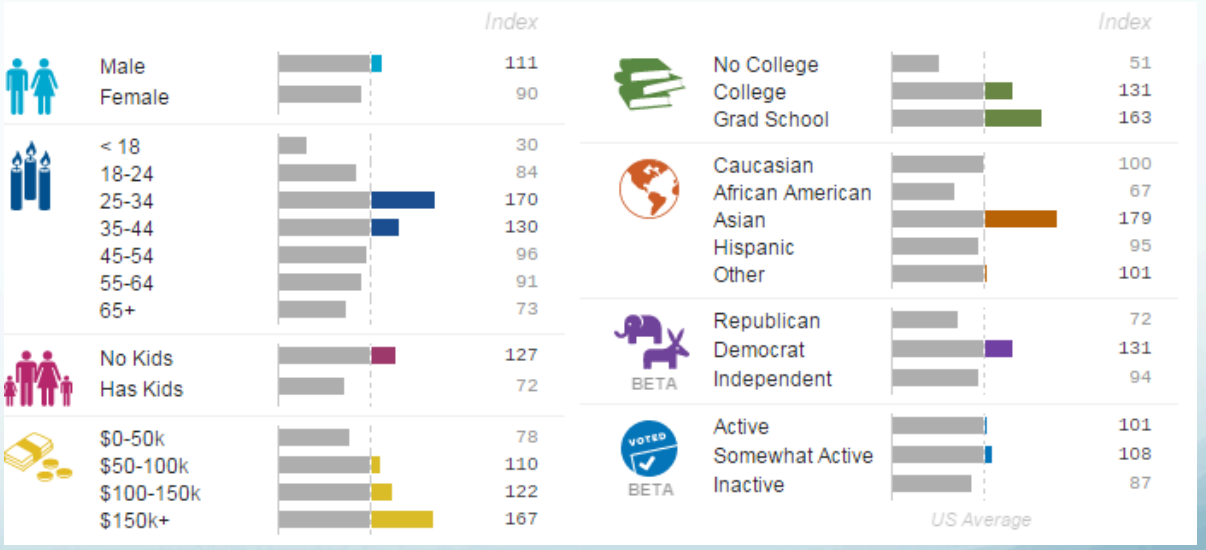
- Outstanding Paper
  - *"From Non-Negative to General Operator Cost Partitioning"*
- Outstanding Paper Honorable Mention
  - *"Predicting the Demographics of Twitter Users from Website Traffic Data".* Aron Culotta, Nirmal Kumar Ravi and Jennifer Cutler , Illinois Institute of Technology
- Outstanding Student Paper
  - *"Surpassing Human-Level Face Verification Performance on LFW with GaussianFace"*

Predicting the Demographics of Twitter Users from Website Traffic Data. [Aron Culotta et al.]

- Create a distantly labeled dataset, instead of using manually labeled dataset

quantcast Track the demographics of visitors of websites

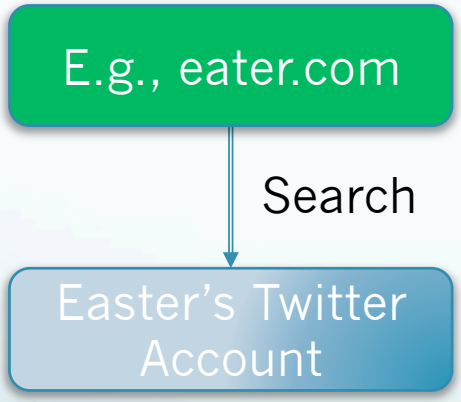
E.g., eater.com

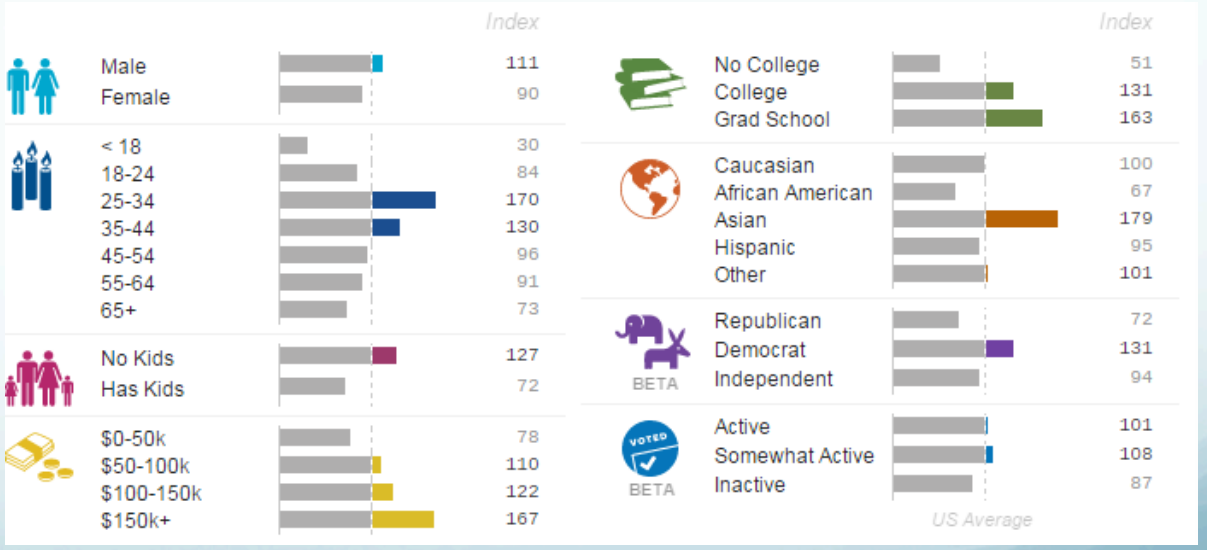


Predicting the Demographics of Twitter Users from Website Traffic Data. [Aron Culotta et al.]

- Create a distantly labeled dataset, instead of using manually labeled dataset

quantcast Track the demographics of visitors of websites

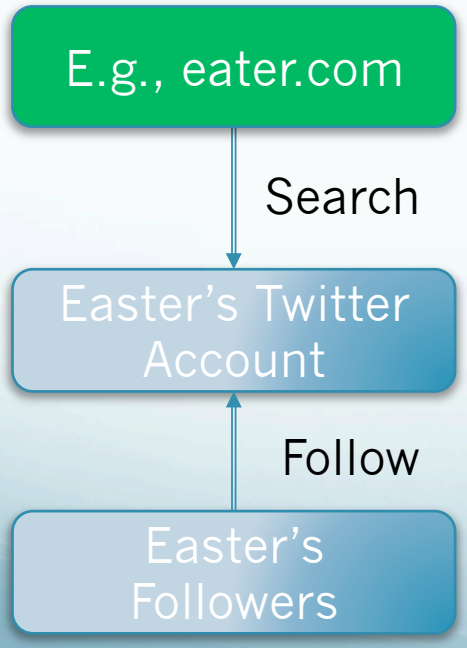


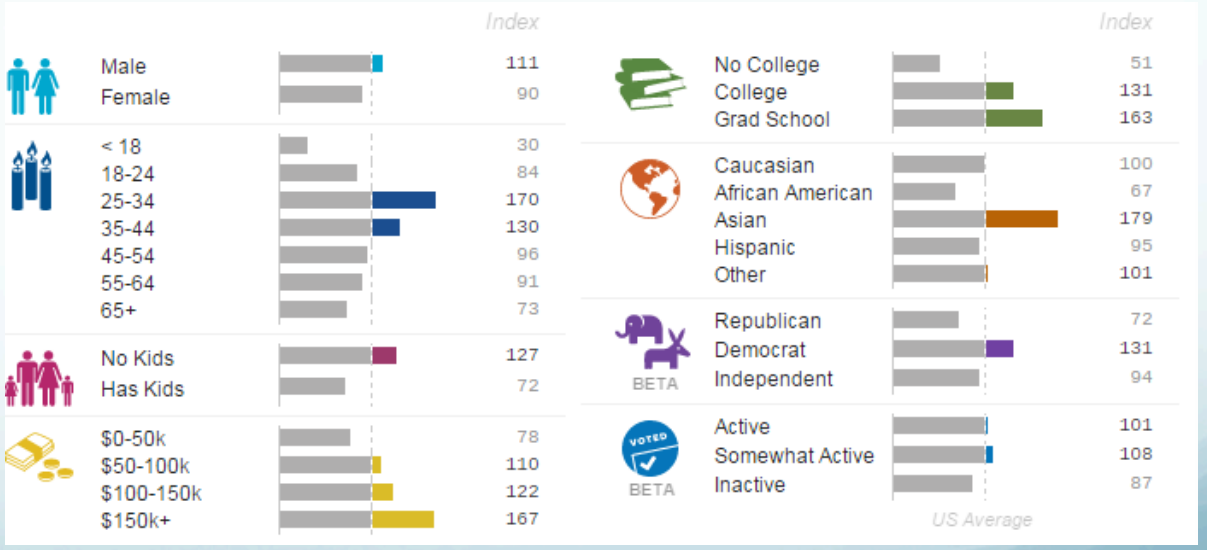


Predicting the Demographics of Twitter Users from Website Traffic Data. [Aron Culotta et al.]

- Create a distantly labeled dataset, instead of using manually labeled dataset

quantcast Track the demographics of visitors of websites

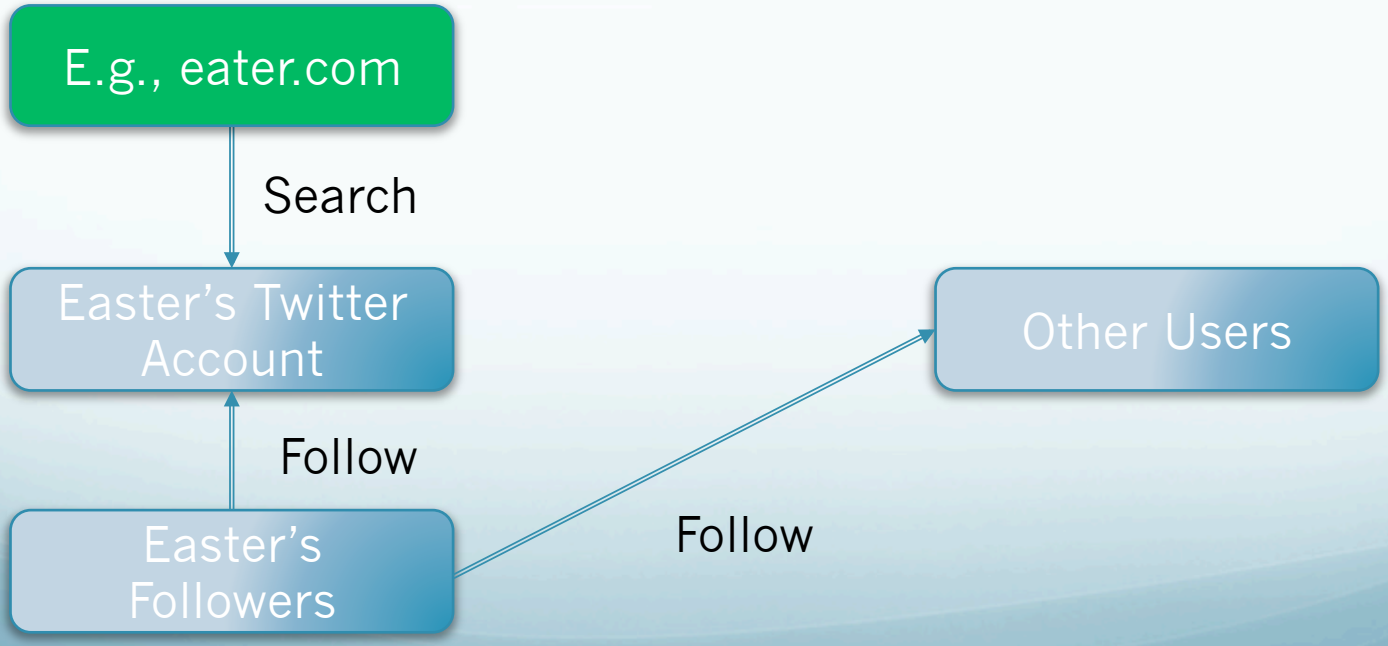




Predicting the Demographics of Twitter Users from Website Traffic Data. [Aron Culotta et al.]

- Create a distantly labeled dataset, instead of using manually labeled dataset

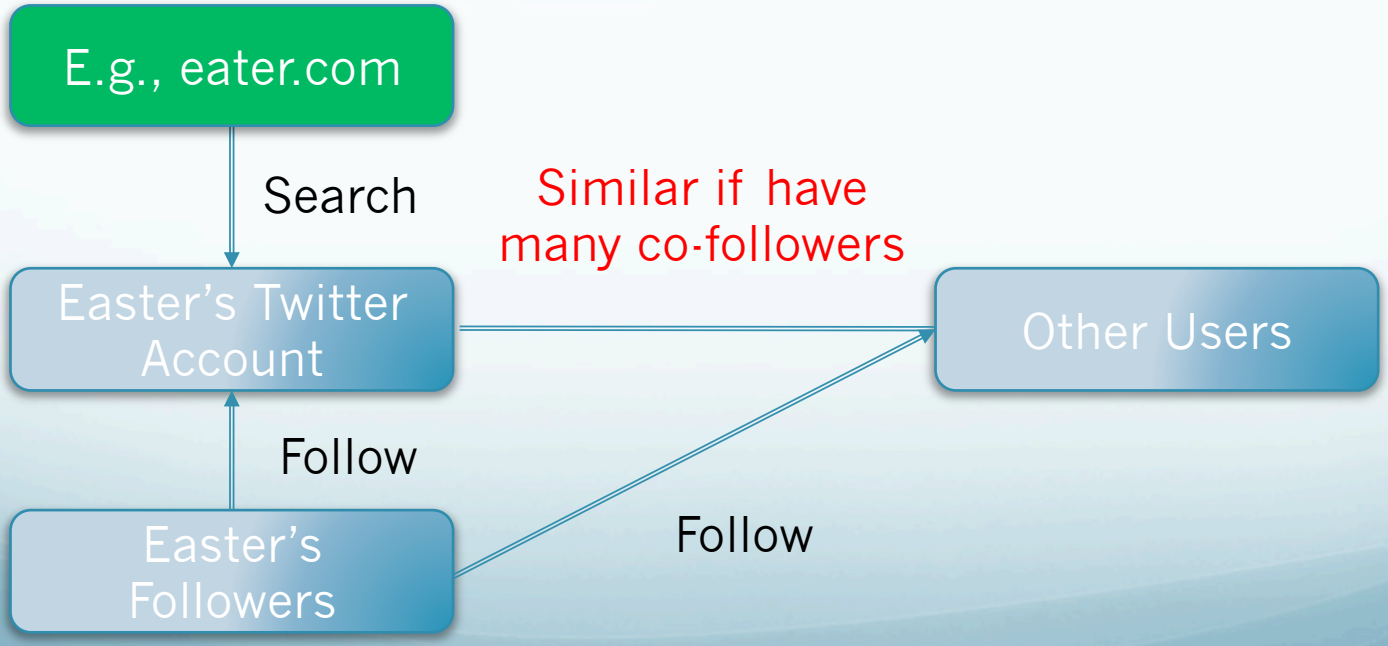
quantcast Track the demographics of visitors of websites



Predicting the Demographics of Twitter Users from Website Traffic Data. [Aron Culotta et al.]

- Create a distantly labeled dataset, instead of using manually labeled dataset

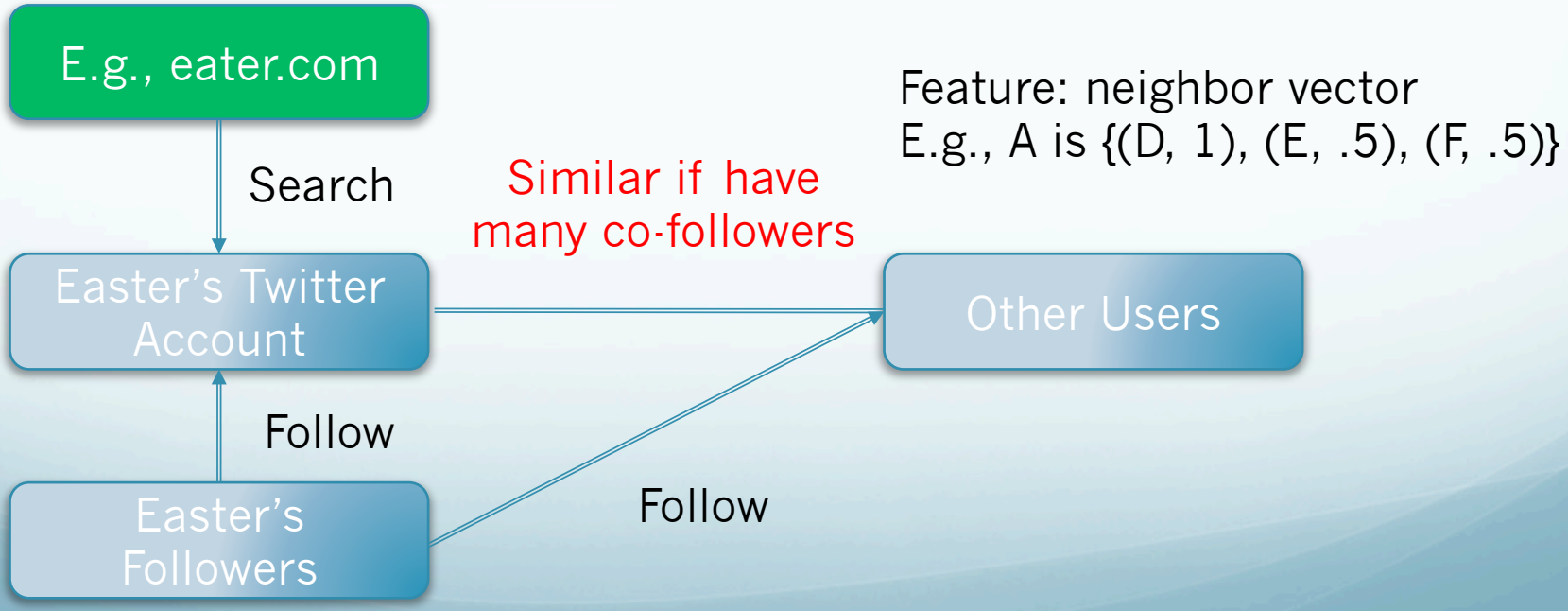
quαntcast Track the demographics of visitors of websites



Predicting the Demographics of Twitter Users from Website Traffic Data. [Aron Culotta et al.]

- Create a distantly labeled dataset, instead of using manually labeled dataset

quantcast Track the demographics of visitors of websites



Predicting the Demographics of Twitter Users from Website Traffic Data. [Aron Culotta et al.]

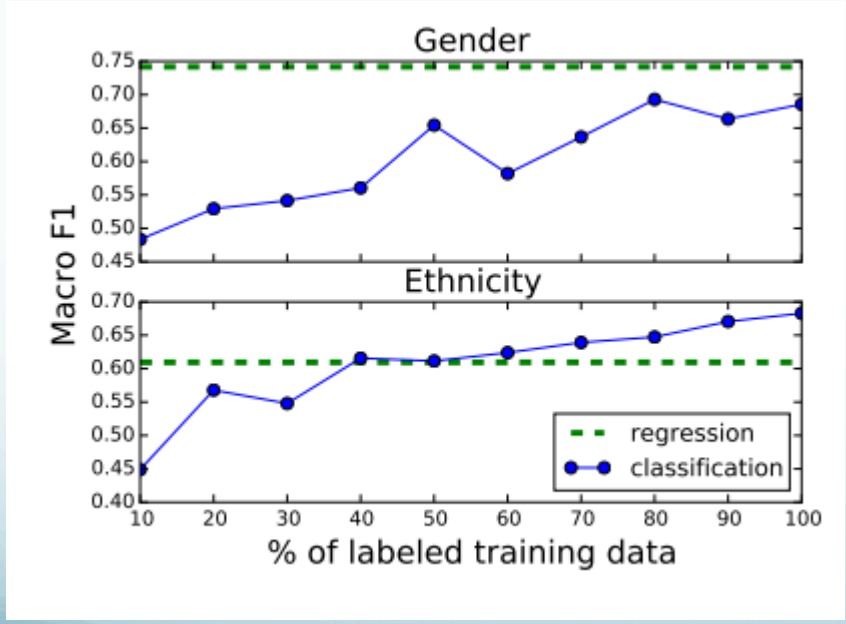
- 6 variables: gender, age, income, education, children, ethnicity
- Regression using both L1 and L2 regularizer

$$\beta^* \leftarrow \text{argmin}_\beta \sum_{j=1}^{M} \frac{1}{N_j} \sum_{i=1}^{N_j} (y_i^{(j)} - \beta^{(j)T} \mathbf{x}_i^{(j)})^2 + \lambda_1 \sum_{k=1}^{p} ||\beta_k||_1 + \lambda_2 ||\beta||_2^2$$

- Evaluation 1: correlation coefficient between the predicted and true demographic variables
  - E.g., predict 30% is female, and quantcase says 40% is female
  - Overall correlation is very strong: 0.77 on average

Predicting the Demographics of Twitter Users from Website Traffic Data. [Aron Culotta et al.]

- Evaluation 2: Macro-F1 for ethnicity and gender
- Manually labeled 615 users and trained a logistic regression classifier



Predicting the Demographics of Twitter Users from Website Traffic Data. [Aron Culotta et al.]

- Evaluation 2: Macro-F1 for ethnicity and gender
- Manually labeled 615 users and trained a logistic regression classifier