

ERCIM SNA Group

Pisa Meeting



Derek Greene



Clique: Graph & Network Analysis Cluster
School of Computer Science & Informatics
University College Dublin, Ireland



Crowdsourcing



- Take a task traditionally performed by an expert and outsource it to a large group of non-experts.
- *Amazon Mechanical Turk*: "Complete simple tasks that people do better than computers and get paid for it."

The screenshot shows the Amazon Mechanical Turk dashboard. At the top left is the logo "amazonmechanical turk" with "beta" and "Artificial Artificial Intelligence" below it. Navigation tabs include "Your Account", "HITs", and "Qualifications". On the right, it says "Already have an account? Sign in as a Worker | Requester". A central banner reads "Mechanical Turk is a marketplace for work. We give businesses and developers access to an on-demand, scalable workforce. Workers select from thousands of tasks and work whenever it's convenient. 38,290 HITs available. View them now." Below this are two main sections: "Make Money by working on HITs" and "Get Results from Mechanical Turk Workers".

Make Money by working on HITs

HITs - *Human Intelligence Tasks* - are individual tasks that you work on. [Find HITs now.](#)

As a Mechanical Turk Worker you:

- Can work from home
- Choose your own work hours
- Get paid for doing good work

Find an interesting task → Work → Earn money

[Find HITs Now](#)

Get Results from Mechanical Turk Workers

Ask workers to complete HITs - *Human Intelligence Tasks* - and get results using Mechanical Turk. [Get started.](#)

As a Mechanical Turk Requester you:

- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITs completed in minutes
- Pay only when you're satisfied with the results

Fund your account → Load your tasks → Get results

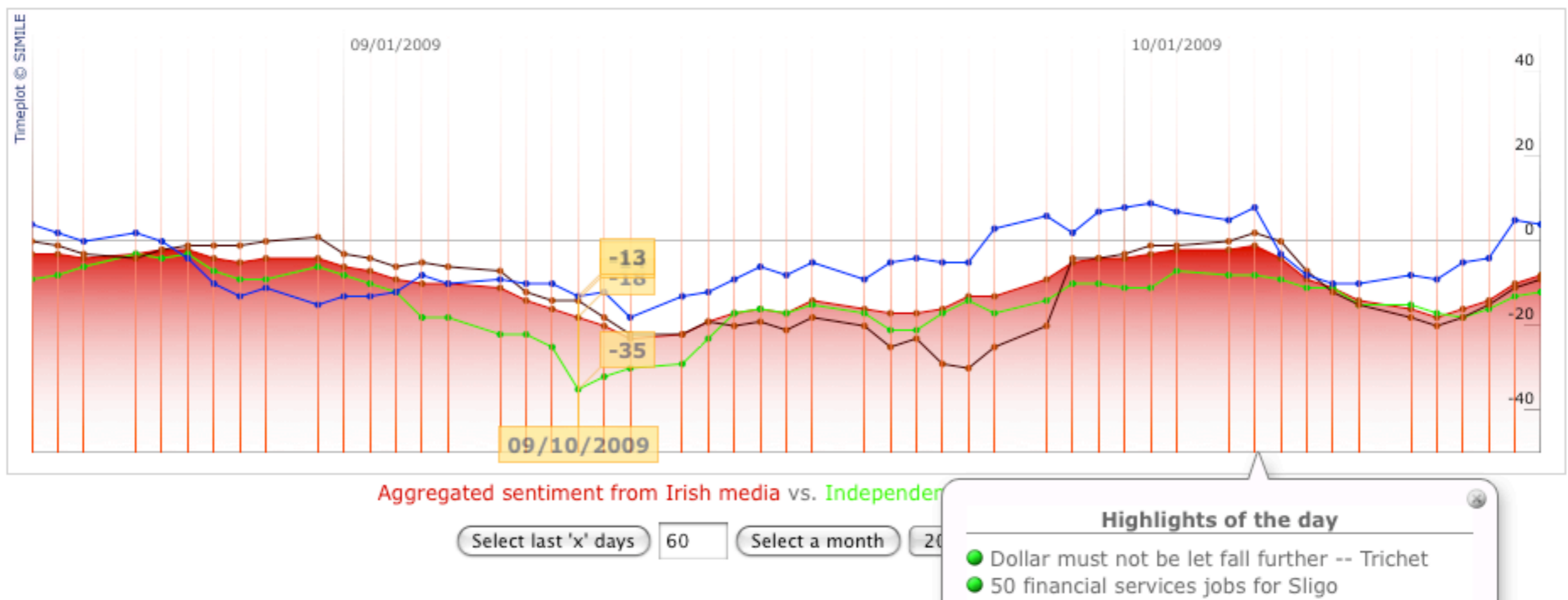
[Get Started](#)

Sentiment Analysis



- System combining *Crowdsourcing + Machine Learning* techniques to track economic sentiment across Irish online media sources (Brew et al, 2009).

<http://sentiment.ucd.ie>



Sentiment Analysis



- Tracked sentiment in 3 news sources over 3 months, collecting over 11,000 articles.

The image shows two overlapping screenshots of news websites. The left screenshot is from IRISHTIMES.com, dated Tuesday, November 3, 2009, with a temperature of 11°C in Dublin. The main article is titled "Facebook to double Irish operation" and discusses the company's plans to increase its workforce in Ireland. A "Related" section lists two other articles: "Networking sites still most popular with Irish web users" and "Facebook to set up international HQ in Dublin". The right screenshot is from RTÉ Business, dated Tuesday, 3 November 2009 16:55. The main article is titled "Tax take to be €2 billion behind target" and reports that the Department of Finance figures show a deficit of €22.7 billion in the first ten months of the year, with tax take falling short of estimates.

Sentiment Analysis



- Tracked sentiment in 3 news sources over 3 months, collecting over 11,000 articles.



- Crowdsourced annotators: 33 non-experts were asked to annotate a subset of the articles each day...

"What impact will this article have on your personal attitude towards the Irish economic situation?"

Positive

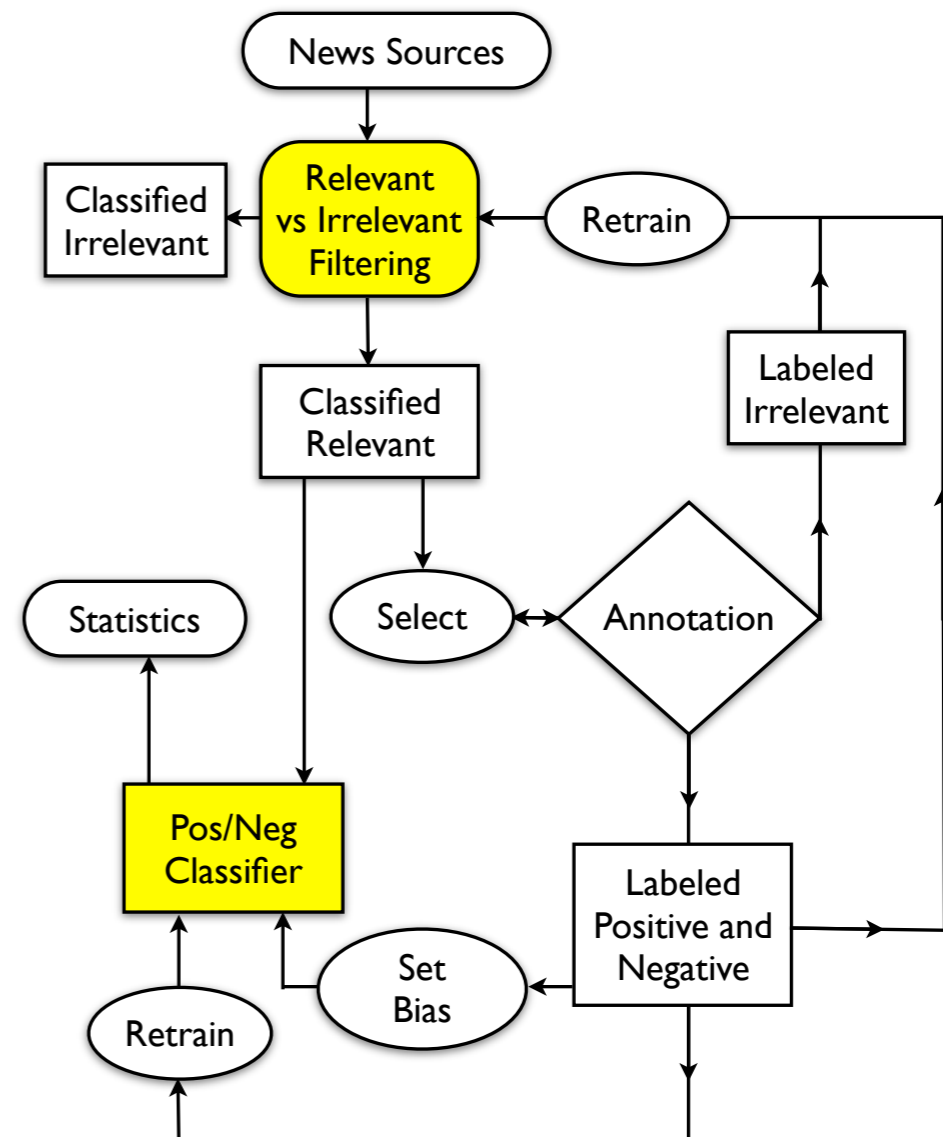
Negative

Irrelevant

System Outline



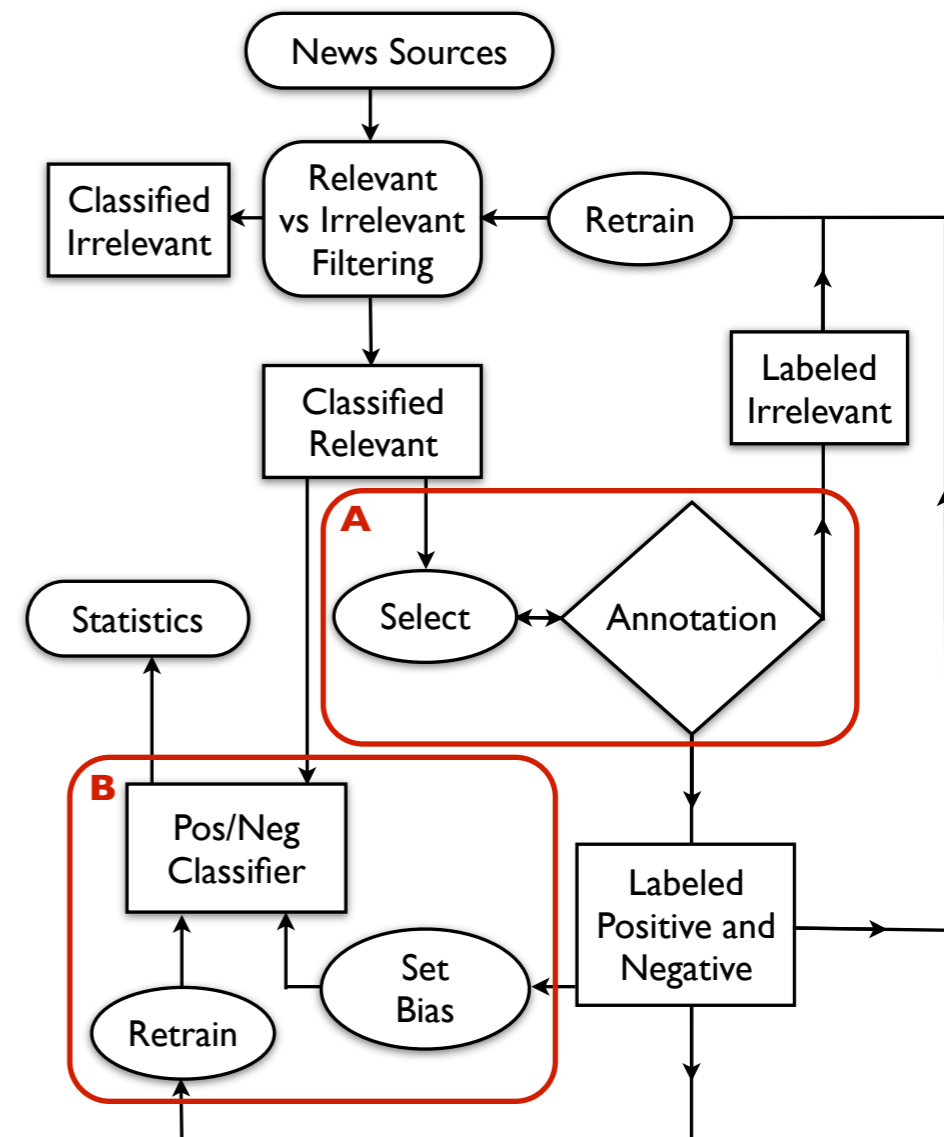
- Two classification problems:
 1. Filter articles irrelevant to economic situation
 2. Positive/Negative (Polarity)



System Outline



- Two classification problems:
 1. Filter articles irrelevant to economic situation
 2. Positive/Negative (Polarity)
- Core components:
 - (A) Selection of subset of articles for annotation
 - (B) Polarity classifier with bias correction



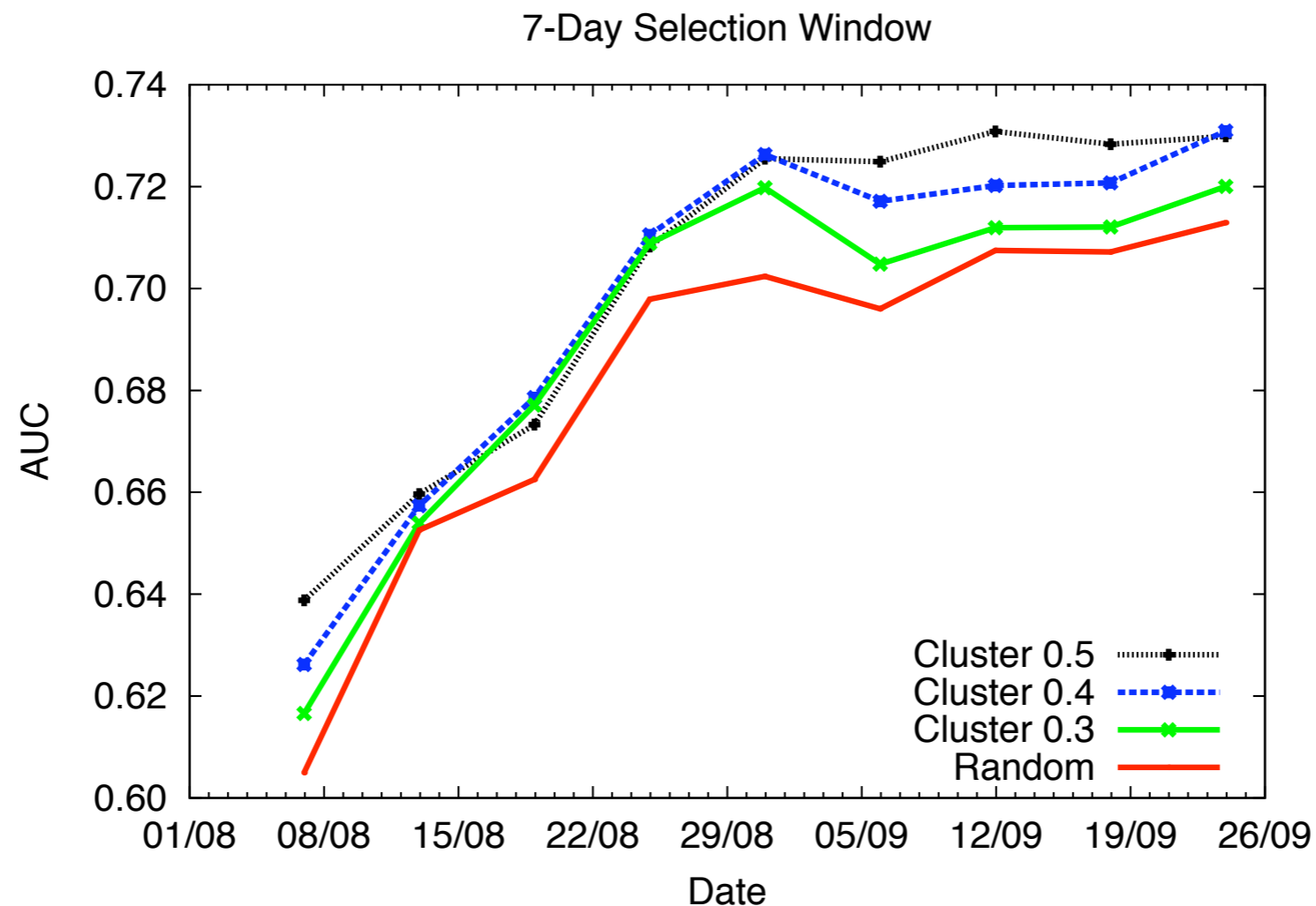
- Q. How do we *select* a useful/interesting set of articles for users to annotate?
- Q. How does *consensus* between users affect the classifier?
- Q. If we have a limited *annotation budget*, how can we make the best use of this budget?
- Q. How can we produce aggregate trends in an *unbiased* way?

- Used a clustering-based strategy to select diverse subset of $\sim 10/300$ articles per day for annotation:
 1. Agglomerate articles on same story.
 2. Weighted farthest-first traversal to choose set of large & diverse clusters.

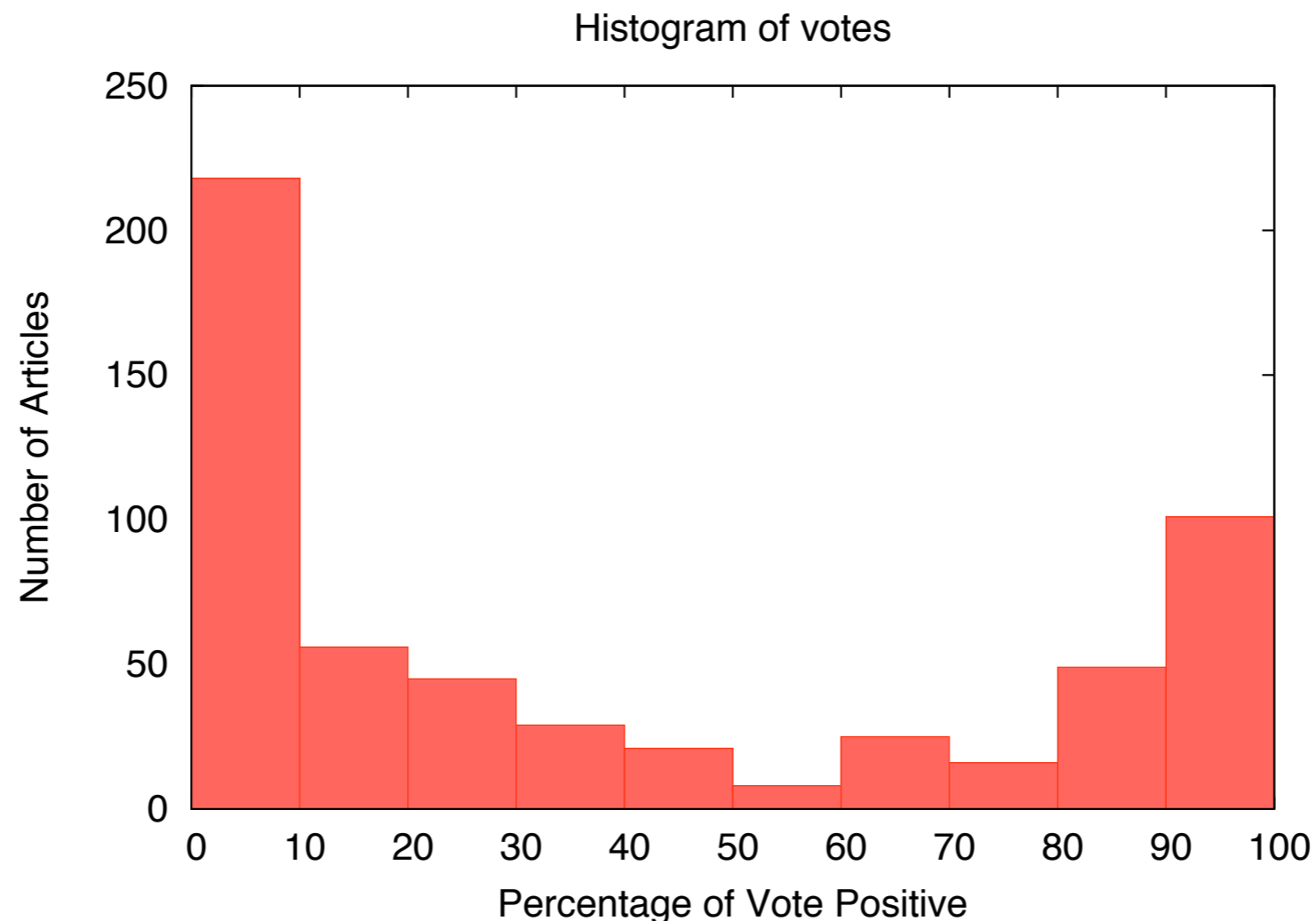
Article Selection



- Used a clustering-based strategy to select diverse subset of $\sim 10/300$ articles per day for annotation:
 1. Agglomerate articles on same story.
 2. Weighted farthest-first traversal to choose set of large & diverse clusters.



- Multiple votes on the same article can be used to derive a majority *consensus* label for an article.
- Level of agreement in the data was surprisingly high: 55% of article consensus decisions are unanimous.



Consensus v. Coverage

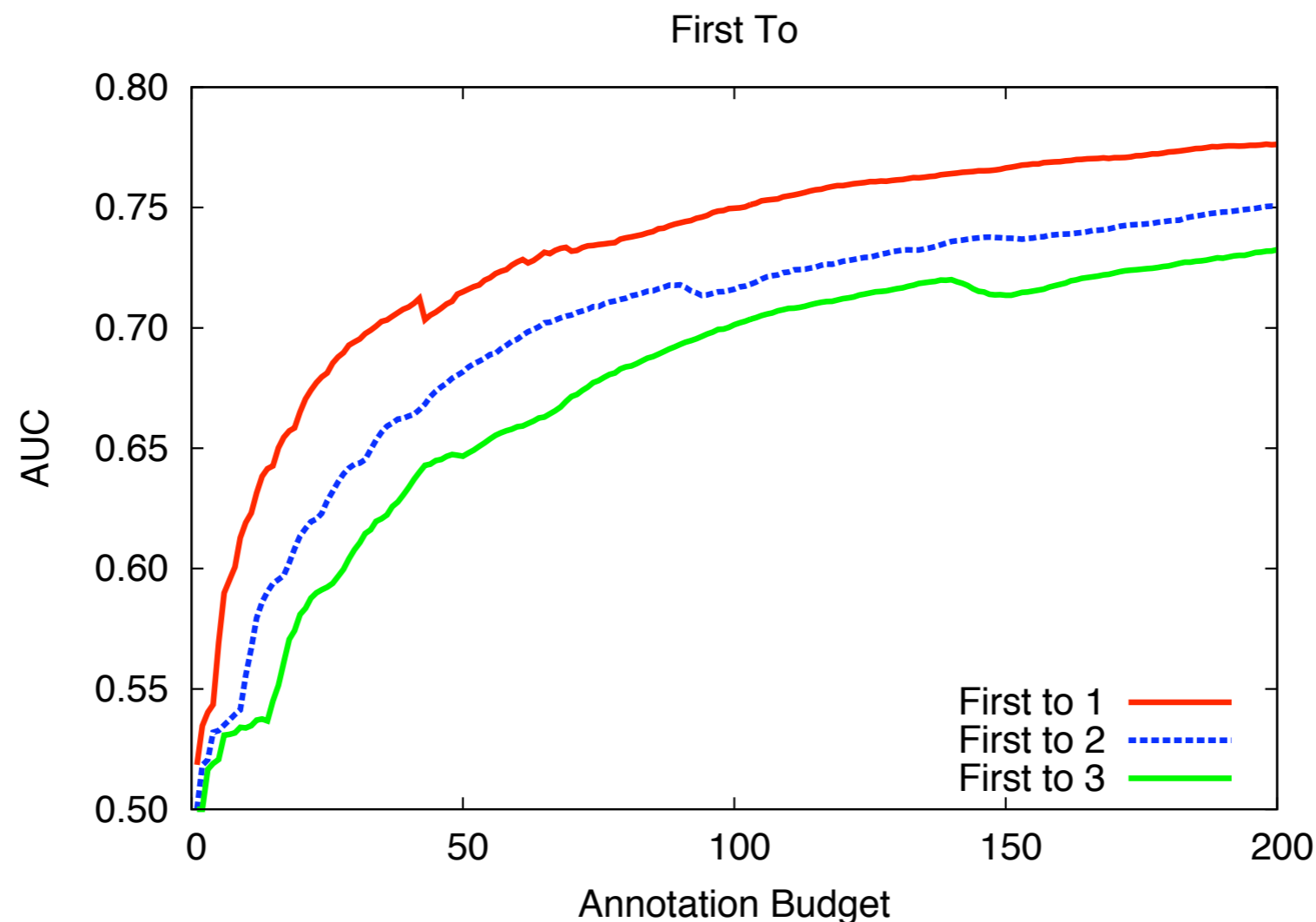


- We would like to cover as much of the domain as possible on a limited annotation budget.
 - ➔ How should we spend that budget?
 - ➔ Do we need multiple annotations per article?

Consensus v. Coverage



- We would like to cover as much of the domain as possible on a limited annotation budget.
 - ➔ How should we spend that budget?
 - ➔ Do we need multiple annotations per article?



Consensus v. Coverage

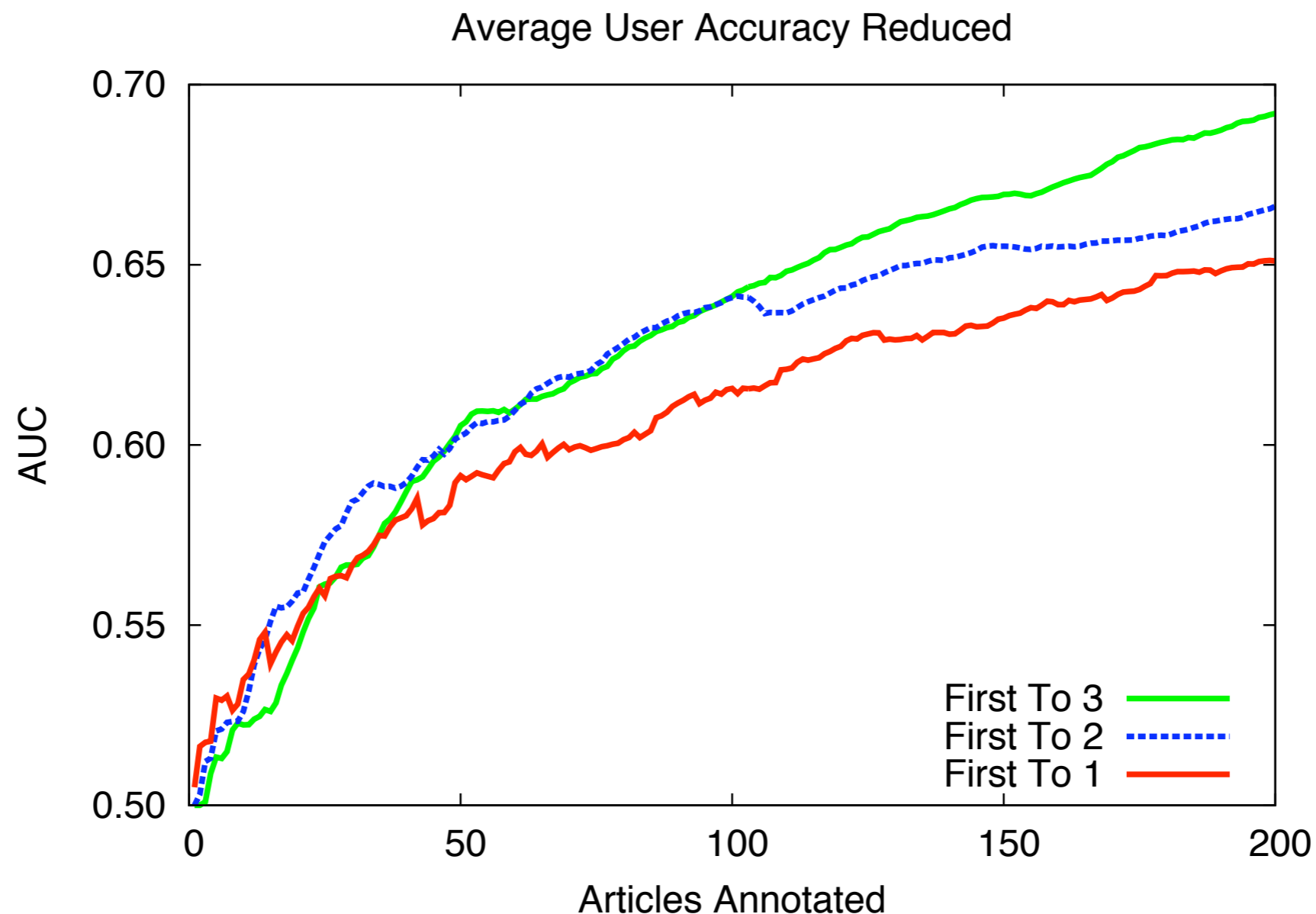


- Level of agreement in our dataset is high.
 - ➔ What happens if we decrease the agreement?

Consensus v. Coverage



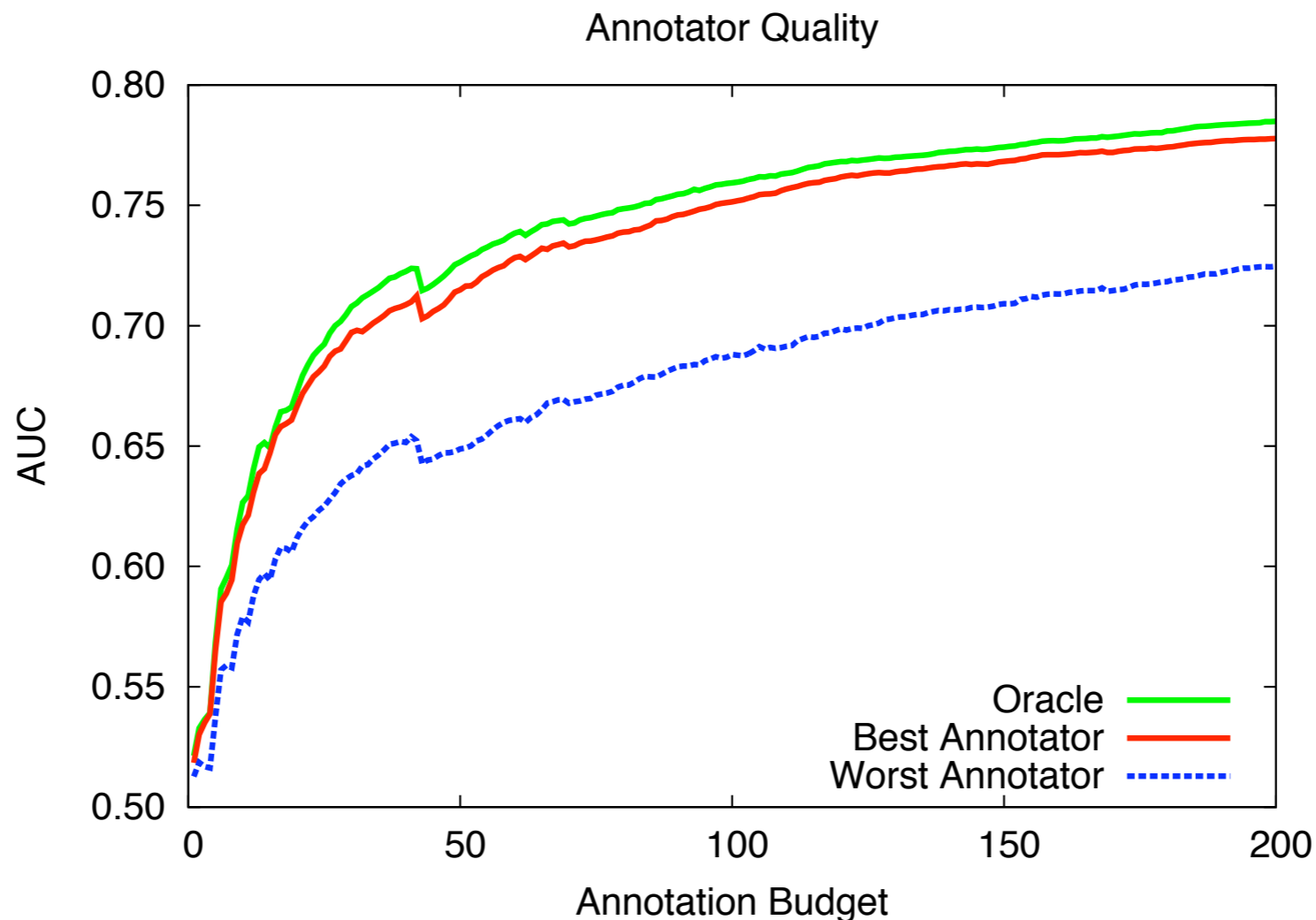
- Level of agreement in our dataset is high.
 - ➔ What happens if we decrease the agreement?



Annotator Quality



- Identifying high-consensus annotators can be effective for making best use of annotation budget.
 - ➔ Lower consensus annotators can be dropped with little deterioration.

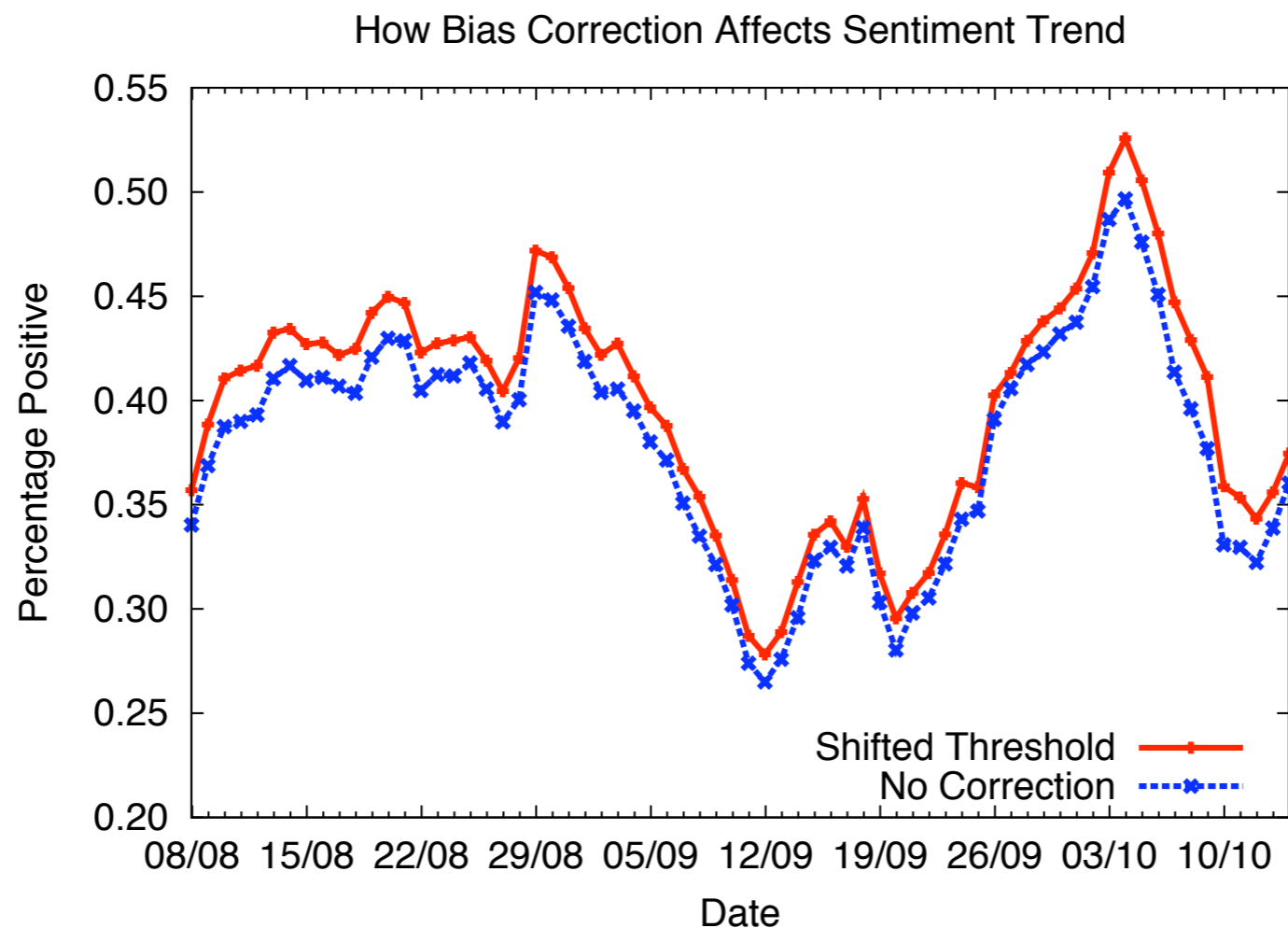


- Classifiers will often become biased towards the dominant class (Bayes, SVMs, k NN...)
- Correct bias by **shifting the threshold** for the Bayes classifier.
 - ➔ Degree of shift is estimated using cross-validation.

Bias Correction



- Classifiers will often become biased towards the dominant class (Bayes, SVMs, *k*NN...)
- Correct bias by **shifting the threshold** for the Bayes classifier.
 - ➔ Degree of shift is estimated using cross-validation.



- Annotators are volunteers - no monetary reward.
- ➔ RSS feed provides summary of day's economic news.
- ➔ Web interface provides interesting summary statistics...

- Annotators are volunteers - no monetary reward.
- ➔ RSS feed provides summary of day's economic news.
- ➔ Web interface provides interesting summary statistics...



- Annotators are volunteers - no monetary reward.
- ➔ RSS feed provides summary of day's economic news.
- ➔ Web interface provides interesting summary statistics...

Top 10 articles ordered by annotations:

[All Time](#) | [Last 7 Days](#)

[Most Positive](#) | [Most Negative](#) | [Most Controversial](#)

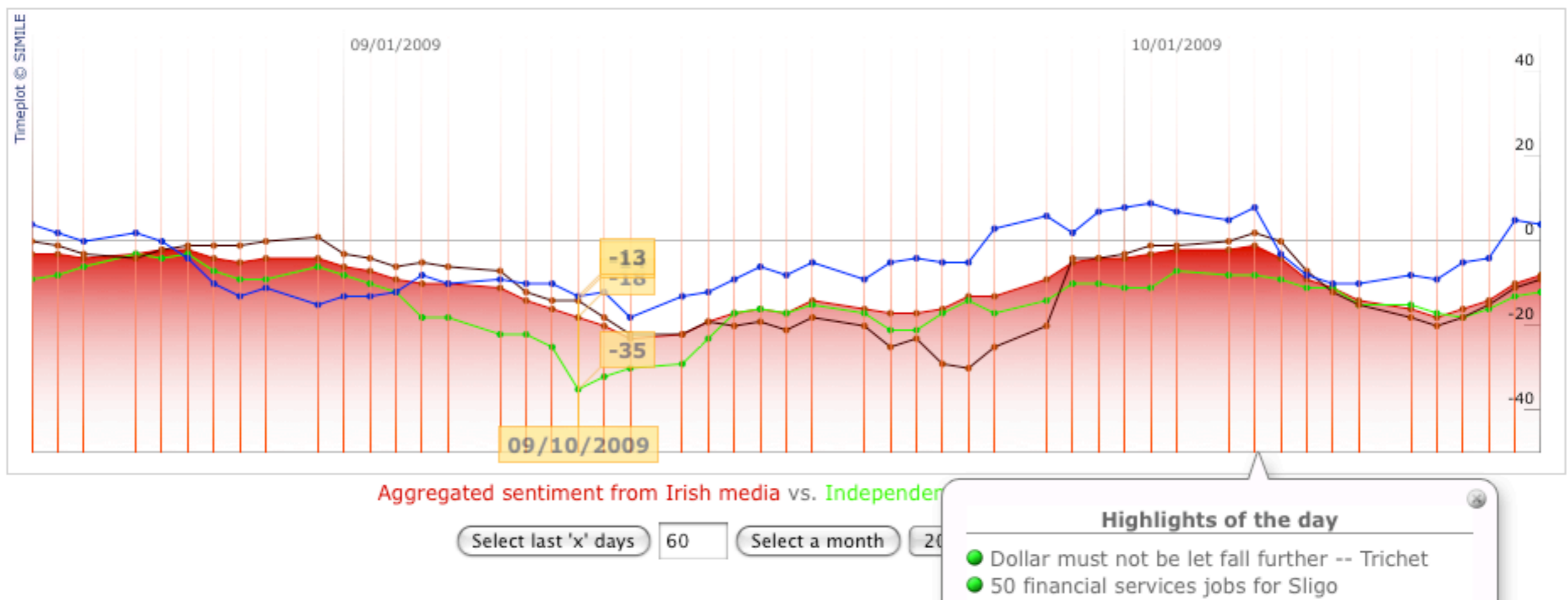
1. [Wireless network 'will create 200 jobs'](#)
2. [Crystal facility to create 100 jobs](#)
3. [PayPal to create 100 jobs at its Dublin centre](#)
4. [IBM raises full-year outlook](#)
5. [Polish president signs Lisbon Treaty](#)
6. [Tullow strikes oil at Ghana well](#)
7. [Biotech firm Amarin in \\$70m funding deal](#)
8. [Abrakebabra group buys O'Briens](#)
9. [Kenny promises to abolish Seanad](#)
10. [Retention of rural transport scheme welcomed](#)

Web Interface



- Annotators are volunteers - no monetary reward.
- ➔ RSS feed provides summary of day's economic news.
- ➔ Web interface provides interesting summary statistics...

Comparison of sentiment across media sources



Any Questions ?

<http://sentiment.ucd.ie>