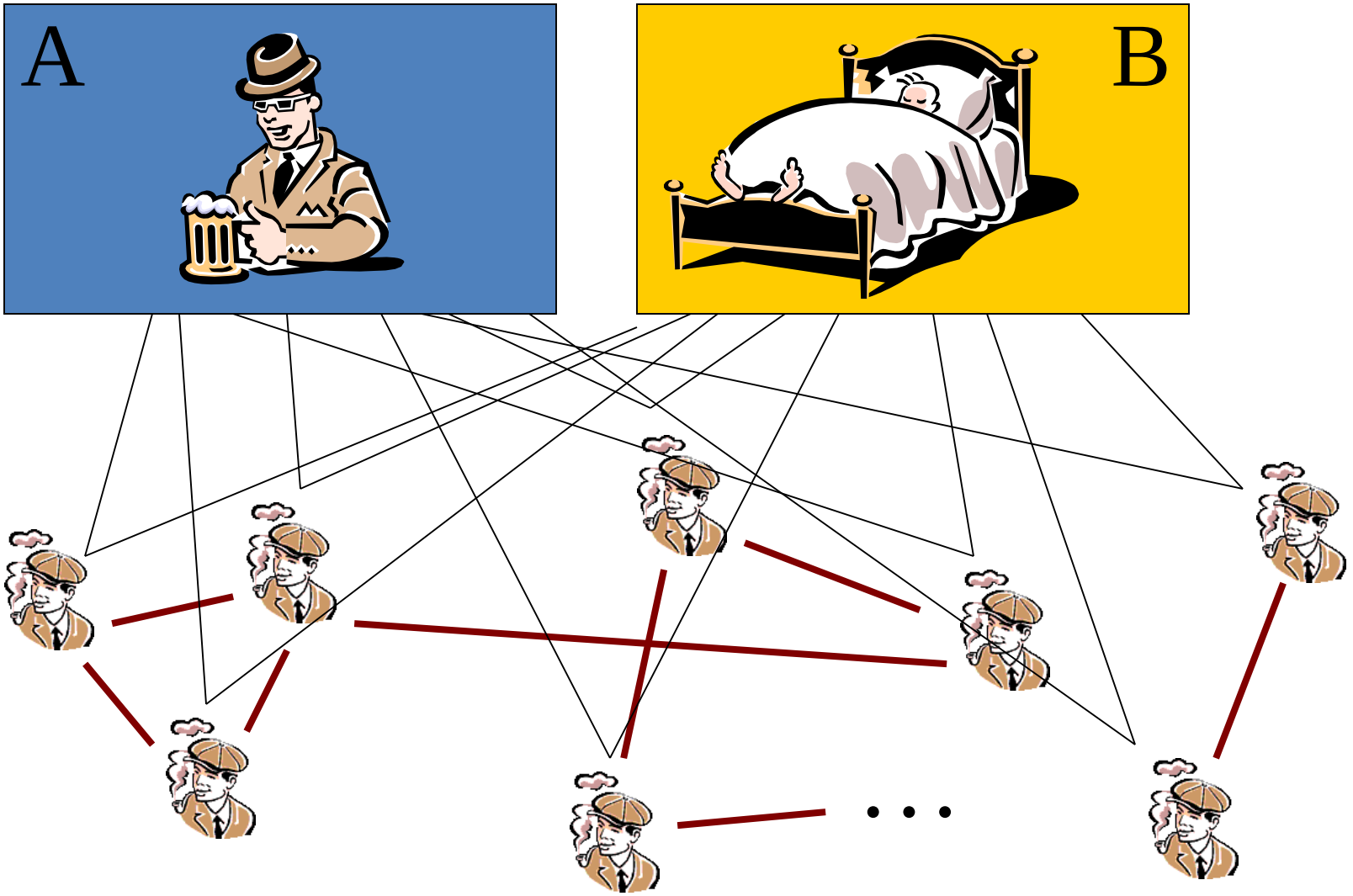
In the search of El-Farol bar problem

"That place is so crowded that no one goes there anymore".



Dancing in the El-Farol!!

- Dancing with the Stars is a dance competition show airing on ABC in the United States, and CTV in Canada.

- The couple receiving the lowest combined total of judges' points and audience votes is eliminated each week until only the champion dance pair remains.

Quantitative Model of the Dance

$$s_i = \lambda q_i + (1-\lambda) v_i$$

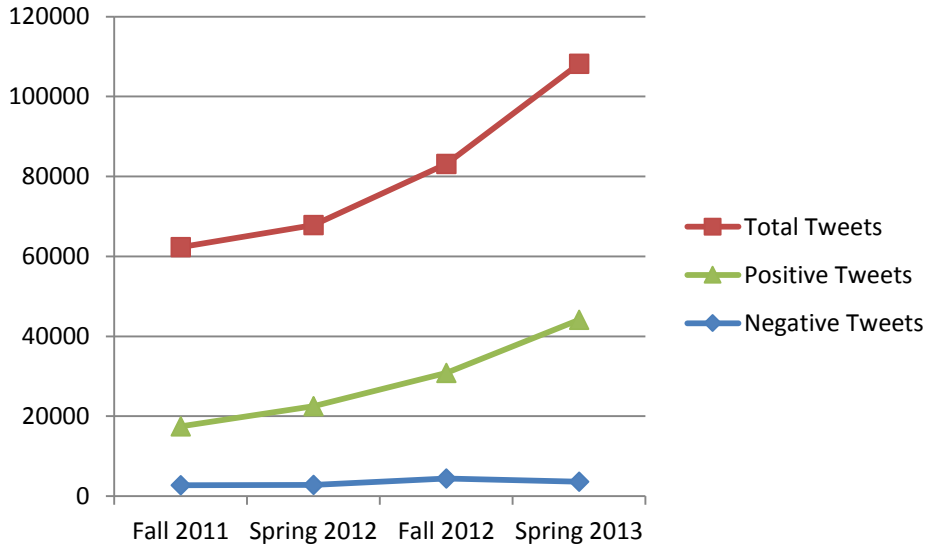
- There is just positive voting, i.e $\lambda < 1$
- According to Wikipedia $\lambda = 1/2$
- The S is not completely known.
- We know who has got eliminated + who is the bottom two of the ranking + the finale ranks
- Assumption: vote is an increasing function of Positive Tweets. $v = f(tweets)$

Data Collection using Topsy

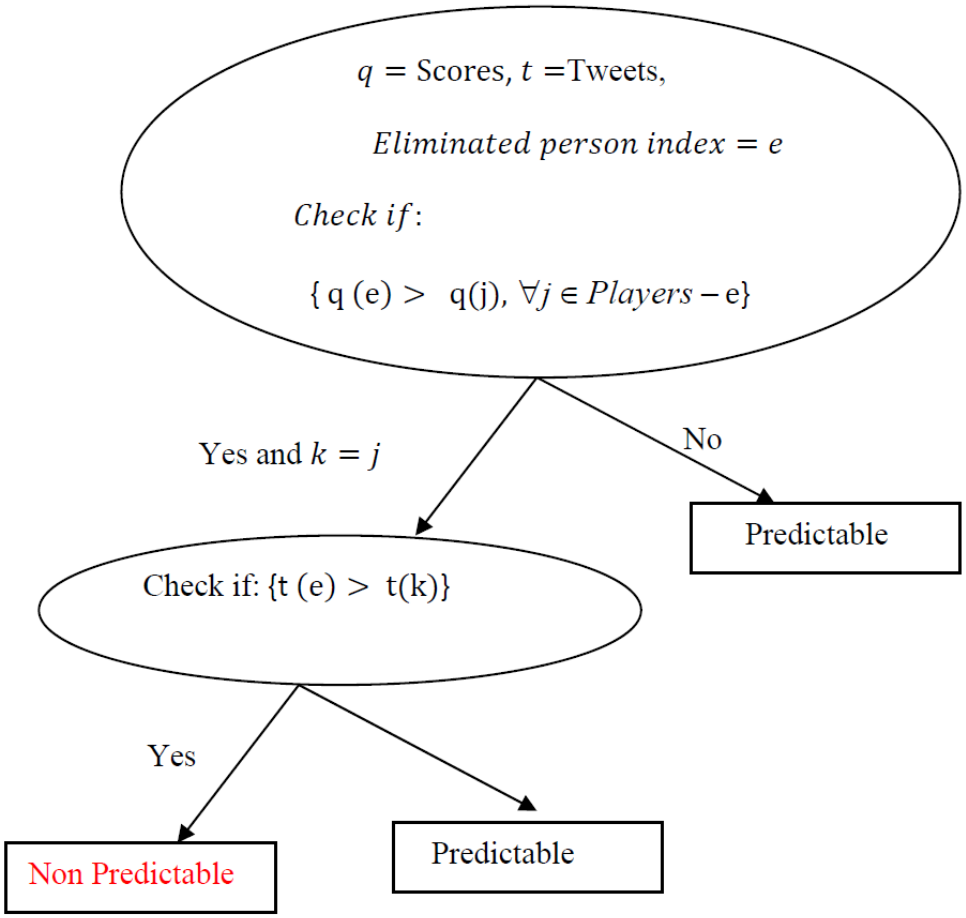
- Difficulties
  - Topsy limitation of data representation + Trick
  - Weekly specific rules, age demographics of twitter users, limited number of the competitions

Data Acquisition

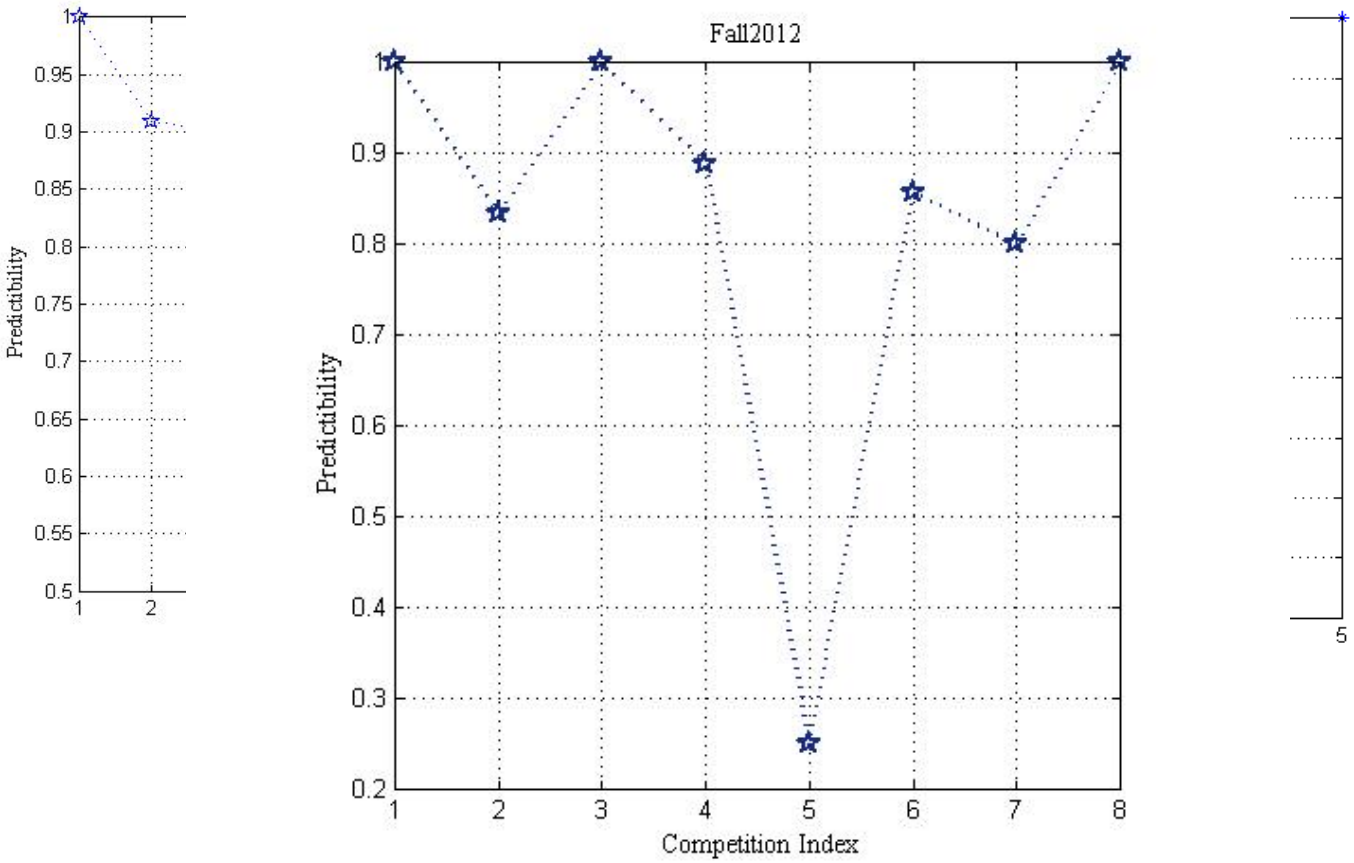
- Working based on tweets and Scores
- Comprehensive database
  - Fall 2011, Spring 2012, Fall 2012, Spring 2013
  - 260 different keyword/Season
  - Python to collect and clean data
- 260,000 tweets



El-Farol observation design

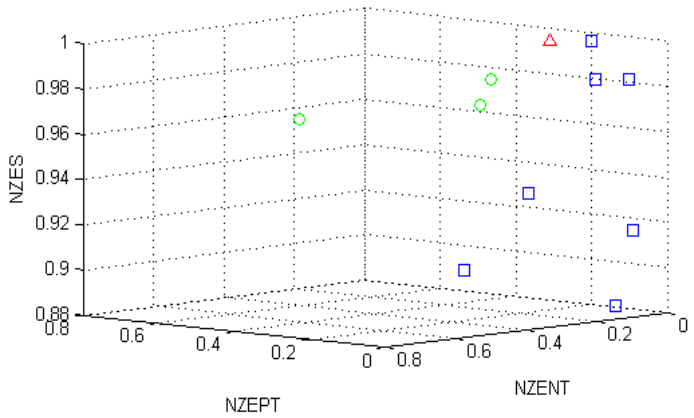


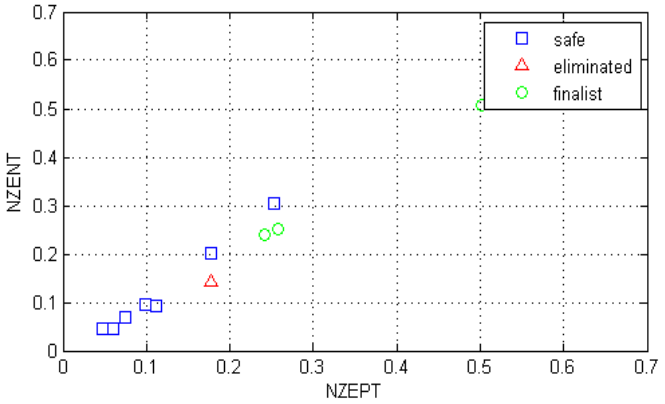
Results

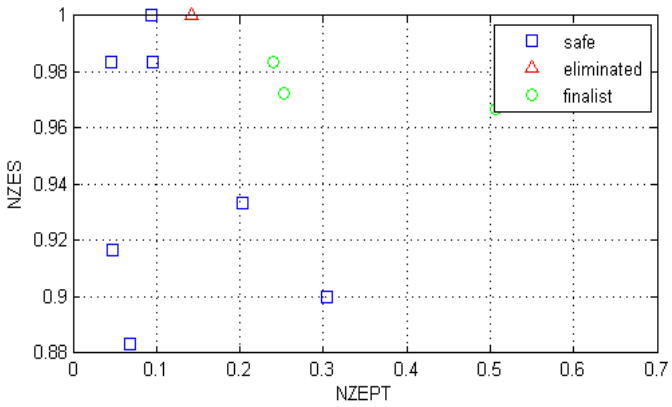

Fall2012
Predictibility
Competition Index
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
2
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
1
2
3
4
5
6
7
8
5

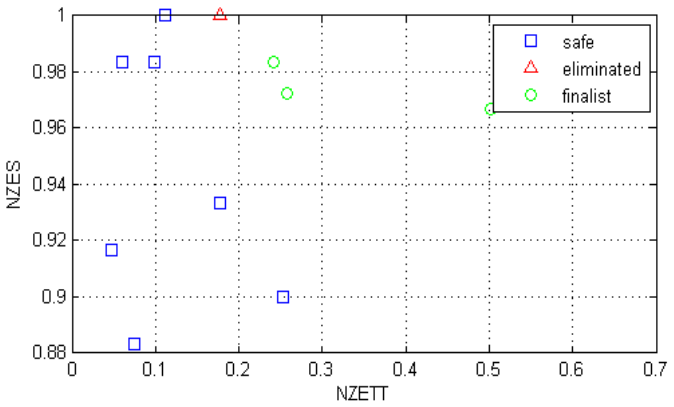
Results

- Elimination 5







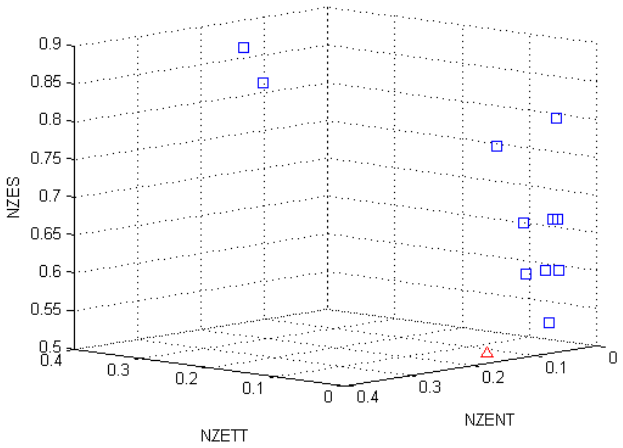


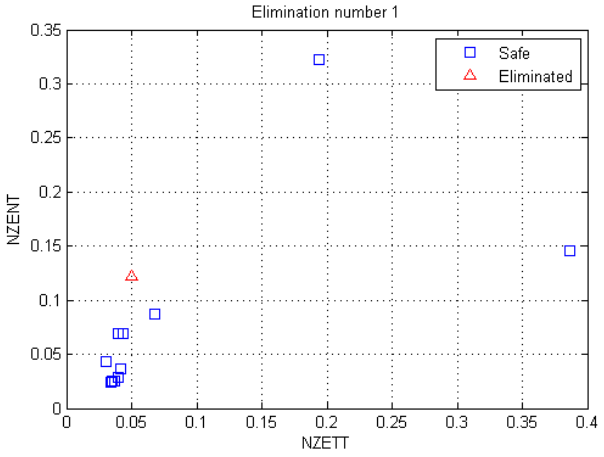
Hypothesis

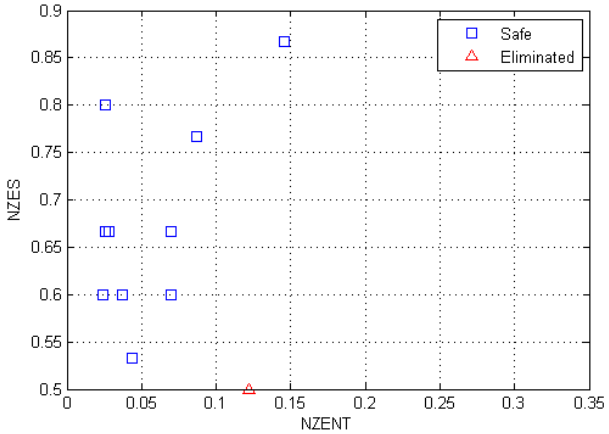
- Hypothesis:
  - Vote is a linear combination of tweets
- $f = \alpha_S S + \alpha_T T + \alpha_N N$
  - f: Ranking function (result)
  - S:score, T: Total tweets, N: Negative tweets
- If so, Feasible region
  - $f^s > f^e$
  - $(A^s - A^e)\,\alpha > 0$
- Feasible region Doesn't Exist!

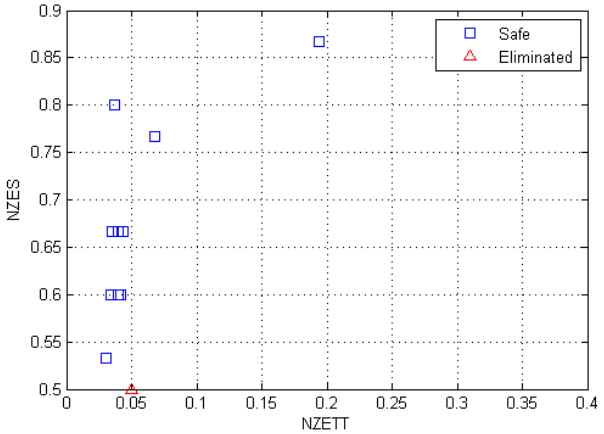
Closer Look at Data

Spring 13



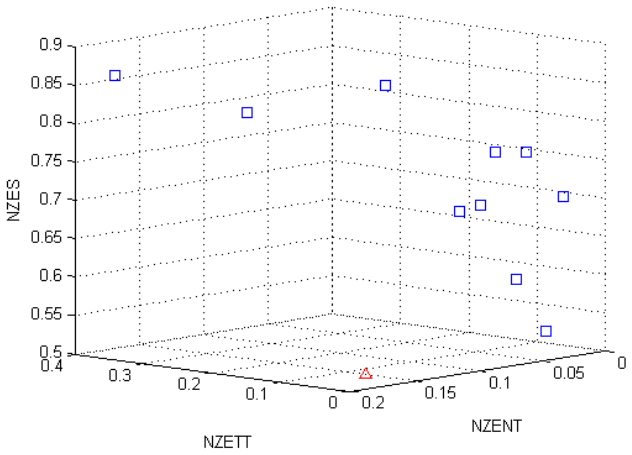


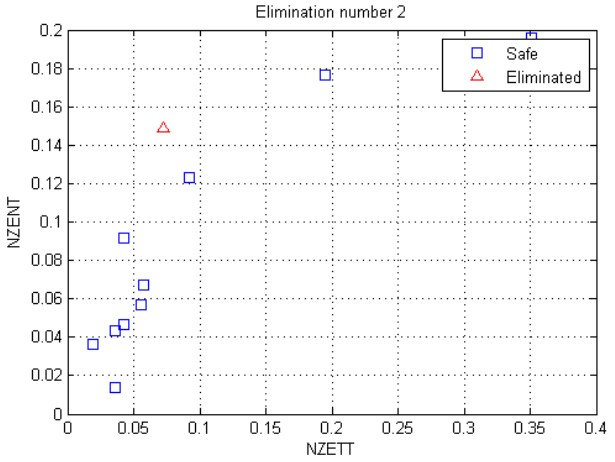


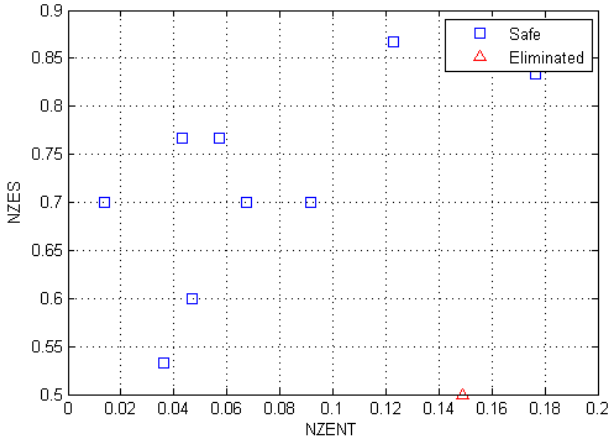


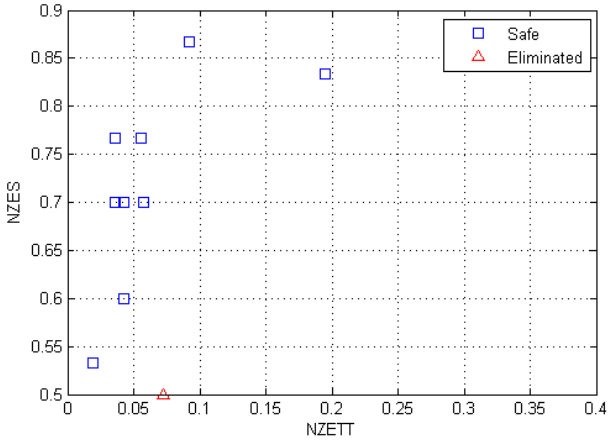
Closer look at data

Spring 13



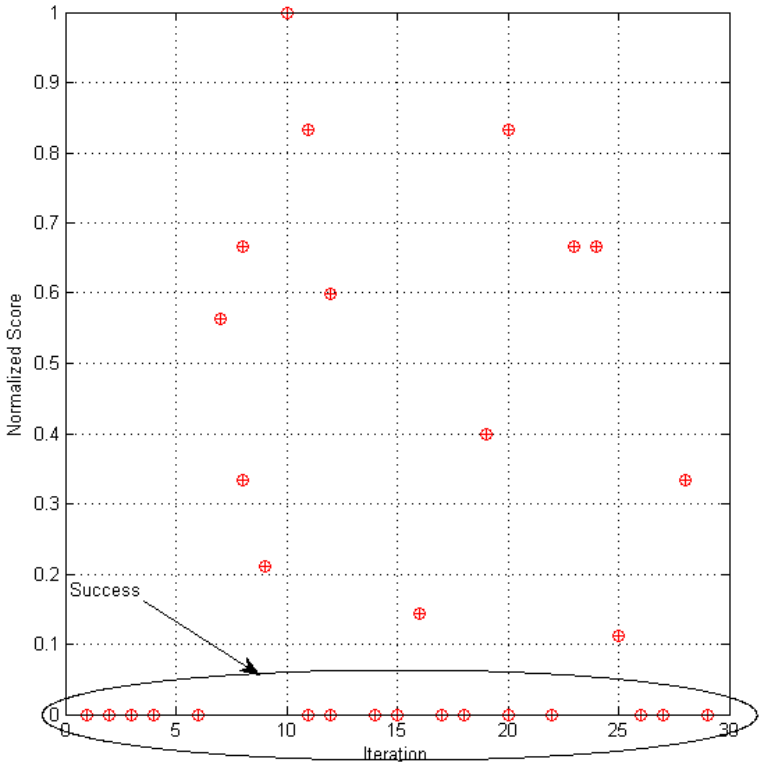






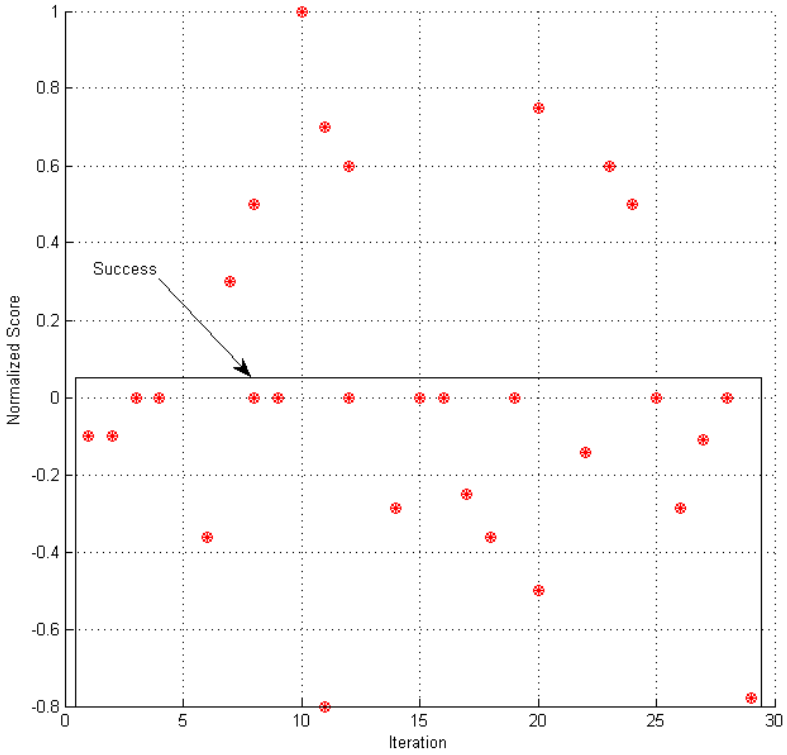
Hypothesis

- Elimination is the Lowest Score!
  - Accuracy = 61%



Hypothesis

- Elimination is form two Lowest Scores!
  - Accuracy = 74%



Conclusion

- El Farol phenomenon
- Linear approximation of tweeter space
- Game of judges