# Will I Reply?

Jonathan Wheeler, Anita Lacea
Stanford University

## Motivation

- Business users on average receive 140 emails and send 43 emails each day [1].
- If only half of these sent emails are replies, then only 15% of the emails users receive require a response.
- **What if our email client could predict which emails need a response before we read them?**

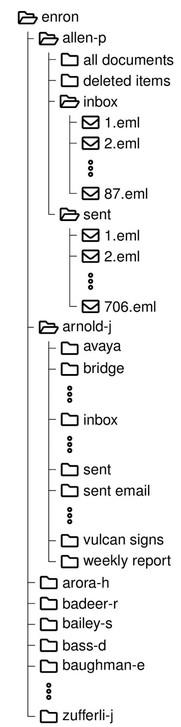
## Data

Personal emails downloaded via Python3 script

- Pros:
  - GMail's file format is standardized across users
  - Emails are tagged with "Reply-To" headers, which make it easy to determine which emails received a response
- Cons:
  - Downloading emails using the script is slow
  - Emails contain personal information
  - Limited to volunteers that are willing to disclose all of their emails to us for this project. This limits sample size

Enron Corpus in public domain

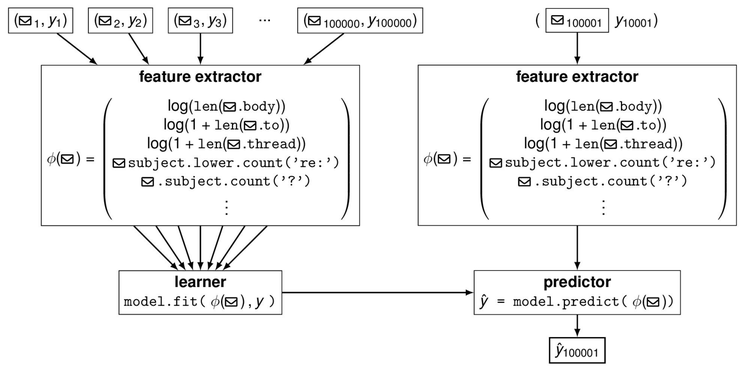
- Pros:
  - Large dataset: contains 150 users and 500,000 emails
- Cons:
  - Emails lack the "Reply-To" header. We have to write a script to search for messages with responses
  - Because email clients and standard practices varied significantly in the early 2000s when this dataset was generated, it is difficult to accurately label the entire dataset algorithmically, and intractable to label manually




Structure of the Enron Corpus on our hard drive

## Model

- Linear Regression
  - Because of the size of the dataset, can be difficult to hold entire dataset in memory
- Stochastic Gradient Descent Regressor
  - Able to handle large datasets
  - Allows successive partial fits, instead of all data at once
    - useful when chunking training set into individual users




## Features

User-independent features (transfer well from user to user):

- (Integer) Number of users in "To:" field
- (Integer) Number of users in "Cc:" field
- (Integer) presence of keywords (e.g. please, RSVP) in body/subject
- (Integer) Length of thread thus far
- (Integer) Length of subject
- (Integer) Length of body)
- (Boolean) Is the recipient in the CC list (and not in the To: list)

User-dependent features (do not transfer well from user to user):

- Does the recipient frequently reply to the sender
- Is the recipients name in the content of the body
- Is the recipient's name in the top line of the body

## Experimental Results

In order to quantify the performance of our experiment, we use the a version of the $F_\beta$ score, which is a weighted mean of *recall* (percentage of true replies captured by predictor) and *precision* (100% - percentage of false positives). We select β = 2 to weight recall twice as important as precision. The marginal cost of having to read an unimportant email is much less than the marginal cost of missing an email requiring a response.

Although the $F_2$ scores in each of the experiments were less than 50% In all of our tests, we found the trained machine learning model to outperform a human in labeling emails.

$$F_\beta = (1+\beta^2)\cdot\frac{\text{precision}\cdot\text{recall}}{(\beta^2\cdot\text{precision})+\text{recall}}$$

Experimental results when run on Enron Corpus

| Model | $F_2$ |
|---|---|
| Linear | 31.7% |
| SGD | 31.5% |
| Human | 12.8% |

## Discussion

- While our algorithm performed better than our human oracle test, it failed to perform at a suitable level to rely on this algorithm in a business setting
- Investigation into other approaches to solving this problem yielded one project that seems particularly relevant: the commercial extension for Google Chrome, "Boomerang for Gmail" [2]
- Both Boomerang and Bromberg & Shultzberg [3] outperform our algorithm. This suggests that user-specific features may be a requirement for the task of determining email reply likelihood
- Future work should consider incorporating combination of user-specific, and multi-user features

## References

[1] S. Radicati, "Email statistics report, 2014-2018." http://www.radicati.com/wp/wp-content/uploads/2014/01/Email-Statistics-Report-2014-2018-Executive-Summary.pdf.
[2] Boomerang For Gmail: Scheduled Sending and Email Reminders https://www.boomeranggmail.com/
[3] Bromberg, A., & Shutzberg, K. (2016). Prediction of User Intent to Reply to Incoming Emails. Retrieved from http://cs229.stanford.edu/proj2013/BrombergShutzberg-PredictionofUserIntenttoReplytoIncomingEmails.pdf

| id | user | folder | filename | message id | date | from | to | cc | subject | body | did reply | reply id |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | allen-p | notes inbox | 36.eml | 12357410.1075855679611.JavaMail.evans@thyme | 975524760 | christi.nicolay@enron.com | phillip.allen@enron.com | NULL | Re: Talking points about California Gas market | Phillip–To the extent that we can give Chair Hoecker our sp... | No | NULL |
| 2 | allen-p | notes inbox | 19.eml | 25849444.1075855679233.JavaMail.evans@thyme | 976723860 | yild@zdemail.zdlists.com | pallen@enron.com | NULL | Y-Life Daily: Bush will almost definitely be prez / Coach K chats | Y-Life Daily Bulletin: December 13, 2000. Note: If your e-mail ... | No | NULL |
| 3 | allen-p | notes inbox | 50.eml | 18640335.1075855713056.JavaMail.evans@thyme | 989499900 | lisa.jacobson@enron.com | lisa.jacobson@enron.com, kevin.mcgowan@enron.com, ... | NULL | RSVP REQUESTED - Emissions Strategy Meeting.... | Due to some of the problems with my email yesterday, I may not have ... | No | NULL |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 28 | allen-p | notes inbox | 16.eml | 10157885.1075855679164.JavaMail.evans@thyme | 976723260 | rebecca.cantrell@enron.com | phillip.allen@enron.com | NULL | Re: | Phillip – Is the value axis on Sheet 2 of the "socialprices" ... | Yes | 1425 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 373 | allen-p | deleted items | 435.eml | 25488440.1075862163573.JavaMail.evans@thyme | 1006296630 | yevgeny.frolov@enron.com | k..allen@enron.com | tim.o'rourke@enron.com, brad.coleman@enron.com, ... | Zero Option | Phillip, It will go along this lines: (Conservative) ... | Yes | 898 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 898 | allen-p | sent items | 3.eml | 20939836.1075855376722.JavaMail.evans@thyme | 1006810558 | k..allen@enron.com | h..lewis@enron.com, dutch.quigley@enron.com, ... | NULL | FW: Zero Option | The project coordinators believe they can reach a solution ... | No | NULL |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 1424 | allen-p | sent | 155.eml | 4811710.1075855683141.JavaMail.evans@thyme | 965397600 | phillip.allen@enron.com | chris.gaskill@enron.com | NULL | empty string | can you build something to look at the historical prices from which ... | No | NULL |
| 1425 | allen-p | sent | 556.eml | 17099882.1075855719049.JavaMail.evans@thyme | 976792500 | phillip.allen@enron.com | jay.reitmeyer@enron.com | NULL | Re: | ———Forwarded by Philip K Allen/HOU/ETC | No | NULL |
| 1426 | allen-p | sent | 468.eml | 16542342.1075855717006.JavaMail.evans@thyme | 981727500 | phillip.allen@enron.com | stagecoachmama@hotmail.com | NULL | empty string | Lucy, Here is a draft of a memo we should distribute to the ... | No | NULL |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

Sample entires in the SQLite3 file we used to index and query our dataset