read.me

Sean Ren, Peter Scheibel, Kha Nguyen, & Michael Mathews
• News in a stream
• News in a stream
• Learn people's preferences on the fly
• News in a stream
• Learn people's preferences on the fly
• Only show stories that people really care about
• News in a stream
• Learn people's preferences on the fly
• Only show stories that people really care about

Reading news has never been that easy!
Technology Stack

Technical
• Framework: Django
• Database: PostgreSQL, SQLite
• Server: Apache
• Front-end: HTML, CSS3, Javascript/jQuery
• 3rd Party API: FeedParser, Beautiful Soup, Readability

Conceptual
• Classification: Naive Bayes Theorem
System Architecture

Server-side

Classifier

DB

Django

Fetch Feeds

Fetch Article

FeedParser

HtmlParser

World Wide Web

Client-side

Javascript

UI
How it works

- Users create categories and associate feeds with categories.
How Classifier works

- Each category has an associated classifier
- Classifier is Naive Bayes with two classes (relevant vs. irrelevant)
- Classifier initially inactive (user provides training examples as they read through articles)
- Once the classifier is active, we decided we always want to return something
- How to do this if all articles are classified as irrelevant?
How Classifier works

- Naive Bayes assigns a score to each contending class, and assigns the classification with the highest score.

- Instead of returning only documents for which \( S_R > S_{-R} \), assign \( S_R - S_{-R} \) to each document.

- Return the documents with the most positive difference.
Recomm-engine Performance

- Tried three variants of classifier
  - Frequency based feature pruning
  - Mutual information based feature pruning
  - No feature pruning
- Question to answer: can feature pruning be used to improve the precision and recall?
As it turns out, using pruning does not improve precision/recall.

We didn't use pruning in the live service.

Although pruning reduces space so it may still be attractive.
User Interface

- Inspired by Flipboard for iPad, New York Times Skimmer
- Mimic newspaper style
- Track user's behavior when reading articles (minimize user interactions)
- implemented with HTML, CSS3, jQuery
Demo

Read.me
Isn't that incredible?
Usability Testing

• See how people use read.me in real world scenarios
• 3 people first round
• Give them tasks, observe how they perform the tasks
• Found lots of bugs and suggestions
• 2 people second round
• Improvements!
What we learned in the usability tests?
Adding feeds is a difficult task

RSS feeds tutorial

1. Find RSS feeds

- Look for an orange (or sometimes, red) RSS icon.
- If a website does not show an RSS icon like this like, find a link called RSS like below.

Weather forecast

<table>
<thead>
<tr>
<th>NewsPulse</th>
<th>U.S.</th>
<th>World</th>
<th>Politics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tools &amp; widgets</td>
<td>RSS</td>
<td>Podcast</td>
<td></td>
</tr>
</tbody>
</table>
Dolphins send Jets to second straight loss, 10-6

*The Seattle Times: Home*

The Miami Dolphins turned two early turnovers into their only points, beating the flat New York Jets 10-6 Sunday at the rainy Meadowlands.
Need feedback when adding a feed to a category (x3)
Bugs found

• Error parsing some feeds
• Hard to get the right content from HTML

How to fix?
• use Beautiful Soup (Python HTML parser)
What we experienced

• Create a web app
• Deploy a Django app on Apache server
• Design an efficient database
• UI design and implementation (cool CSS3 properties)
• Avoid reimplementing code (3rd-party code)
• Classification: NB
• Usability Test
http://readme.cs.washington.edu