Abstract

With machine learning and natural language processing, we're building open-source tools to improve the human review process of crisis reports.

What is Crisis Mapping?

In crisis situations, there's an information gap:

- Contested elections
- Natural disasters
- Humanitarian abuses

On Ushahidi, gathering, reviewing, and mapping incident reports to narrow this gap.

The Problem

- Currently, report processing is entirely manual. Each message is individually reviewed: anonymized, categorized, and geo-located.
- Work is slow and tedious. It's difficult to scale up for large volumes or fast-paced situations.

Our Targets

After discussions with the crisis crowdsourcing community, we identified the most pressing needs – automatic suggestions for:

- Classifying report categories
- Flagging private info
- Flagging locations and URLs
- Detecting language
- Detecting (near-) duplicate reports

Methodology & Results

Classifying report categories

- SVMs trained on 1000s of manually-reviewed reports from past crisis maps - overall and per situation.
- Text classification features included a bag-of-words unigram frequency feature set (frequency cut-off = 5).
- Unigram TF-IDF and bigram features didn't help.

The Cold Start Problem

Classifier performance as a function of # training reports

**Lebanon 2009 election**

- **BLUE:** Classifier trained only on the particular election's reports, with improvement over time as more reports are included in training.
- **RED:** Classifier trained on reports from other elections' reports.

**Bottom line:** A classifier pre-trained on other election reports can help solve the cold-start problem for a new election monitoring.

Flagging private info, locations, & URLs

- Use a pre-trained named entity recognizer (NER) to highlight names and locations.
- Use regular expressions to highlight e-mail addresses, phone/ID numbers, and URLs.

Detecting language

SimHash: Efficiently hash each report to a 64-bit representation. (Near-)duplicates have short distances.

Human Evaluation

We’re currently simulating a crisis experiment on live volunteers, testing 3 versions of the review process:

- Entirely manual, without any suggestions (baseline)
- With our system’s suggestions
- With the correct answers “suggested” (upper bound)

We’ll analyze quantitative changes in reviewer speed and accuracy, and collect qualitative feedback.

Future Work

- Train adaptive classifiers (work in progress)
- Suggest new categories via topic modeling
- Create more pre-trained classifiers (e.g. for disasters)
- Estimate report urgency
- Improve tool performance

Collaborators welcome!
github.com/dssg/ushine-learning

This work was done during the Eric & Wendy Schmidt Data Science for Social Good Summer Fellowship at the University of Chicago.