nouri_2020_mining_crowdsourcing_problems_from_discussion_forums_of_workers

Year

2020

Author(s)

Nouri, Zahra and Wachsmuth, Henning and Engels, Gregor

Title

Mining Crowdsourcing Problems from Discussion Forums of Workers

Venue

COLING

Topic labeling

Manual

Focus

Secondary

Type of contribution

Established approach

Underlying technique

Manual labeling

Topic labeling parameters

\

Label generation

Given the seven determined topics, we first interpreted them based on the 15 highest-probability words in the associated word lists.
For reliability, two authors of this paper did so independently, one of which did not know the literature review results.

Both then sat together to derive a problem label for each topic.
To increase the certainty of a correct interpretation, we additionally inspected the 20 highest-probability reviews of each topic.
This indeed led to a slight label change in one case.

Malfunctioning environment		Workload misestimation				
1 Software errors	**2 Failed completion**	**3 Bad time estimation**	**4 High work effort**	**5 Low payment**	**6 Privacy violations**	**7 Unfair rejection**
break, link, page, one, requester, survey, return, question, amazon, problem, report, get, dead, issue, try	code, survey, break, time, submit, complete, end, timer, get, completion, minutes, waste, error, finish, take	pay, minutes, time, low, long, take, question, one, bubble, page, way, would, even, end, get	writing, write, much, cent, worth, require, penny, paragraph, back, prompt, work, photo, want, single, throw	underpay, pay, writing, screener, bad, screen, unpaid, avoid, requester, research, word, question, per, study, number	tos, require, email, violation, information, ask, site, inquisit, firefox, address, error, personal, file, name, website	reject, requester, rejection, work, work, hit, get, check, email, update, answer, response, attention, reason, one

Table 1: The top 15 words of each of the seven crowdsourcing problems found via topic modeling. The label of each problem (column headings) has been assigned manually.

Motivation

\

Topic modeling

LDA

Topic modeling parameters

- α: 1.0/k
- η: 0.1
- Nr of topics (k): {2, 5, 7}

Nr. of topics

7

Label

Two (one per labeler) independently manually assigned single or multi-word label for each topic

Label selection

No formal methodology, but: “Both [authors] then sat together to derive a problem label for each topic. ”

Label quality evaluation

\

Assessors

2 authors (one of which not knowing the literature review results), sitting together to compare labels.

Domain

Domain (paper): Crowdsourcing
Domain (corpus): Workers’ rights

Problem statement

Crowdsourcing is used in academia and industry to solve tasks that are easy for humans but hard for computers, in natural language processing mostly to annotate data. The quality of annotations is affected by problems in the task design, task operation, and task evaluation that workers face with requesters in crowdsourcing processes.
To learn about the major problems, the paper proposes

- A review where problems known from interviews with workers are collected and organized
- An empirical data analysis where topic modeling is used to mine workers’ complaints from a new English corpus of workers’ forum discussions.

Corpus

Origin: Turkopticon

Nr. of documents: 8610

Details:

- 27,041 Turkopticon reviews from February 2017 to November 2018 are crawled
- Reviews with the tag “not recommended” are selected, since it’s hypothesized that they focus on complaints.

Document

A single Turkopticon review, containing:

- A title
- A link
- Tags
- A review body

Pre-processing

- Tokenization to segment each review into words and other tokens
- Removal of numbers (such as dates, times, and prices)
- Removal of stopwords (such as function words (e.g., “at”, “to”, “which”))
- Removal of punctuation, which were observed overproportially in the reviews
- Lemmatization (to merge different inflections of the same word)
- Removal of the tag “not-recommended” found in all negative reviews
- Removal of low-frequency words, namely those that occur only once in all the analyzed reviews.

```
@inproceedings{nouri_2020_mining_crowdsourcing_problems_from_discussion_forums_
of_workers,
    title = "Mining Crowdsourcing Problems from Discussion Forums of Workers",
    author = "Nouri, Zahra  and
      Wachsmuth, Henning  and
      Engels, Gregor",
    booktitle = "Proceedings of the 28th International Conference on
Computational Linguistics",
    month = dec,
    year = "2020",
    address = "Barcelona, Spain (Online)",
```

```
    publisher = "International Committee on Computational Linguistics",
    url = "https://aclanthology.org/2020.coling-main.551",
    doi = "10.18653/v1/2020.coling-main.551",
    pages = "6264--6276",
    abstract = "Crowdsourcing is used in academia and industry to solve tasks that are easy for humans but hard for computers, in natural language processing mostly to annotate data. The quality of annotations is affected by problems in the task design, task operation, and task evaluation that workers face with requesters in crowdsourcing processes. To learn about the major problems, we provide a short but comprehensive survey based on two complementary studies: (1) a literature review where we collect and organize problems known from interviews with workers, and (2) an empirical data analysis where we use topic modeling to mine workers{'} complaints from a new English corpus of workers{'} forum discussions. While literature covers all process phases, problems in the task evaluation are prevalent, including unfair rejections, late payments, and unjustified blockings of workers. According to the data, however, poor task design in terms of malfunctioning environments, bad workload estimation, and privacy violations seems to bother the workers most. Our findings form the basis for future research on how to improve crowdsourcing processes.",
}
```

#Thesis/Papers/Initial