

Building a dataset for Scientific Paragraph Revision annotated with revision instruction

Léane Jourdan*, Florian Boudin*, Akiko Aizawa*, Nicolas Hernandez*, Richard Dufour*

{firstname.lastname}@univ-nantes.fr

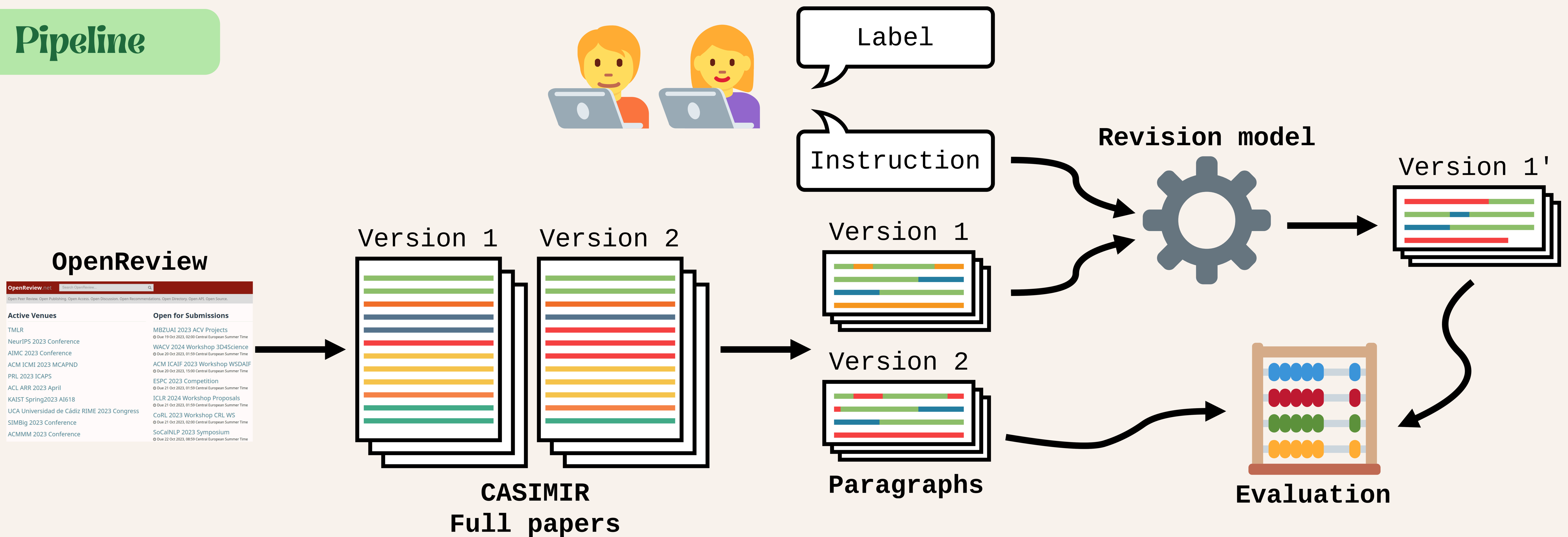
♣ LS2N, Nantes Université ♦ JFLI, NII, Tokyo ♠ NII, Tokyo

Domain: Scientific writing assistance

Task: Paragraph revision

Contribution: Annotated corpus

Pipeline



Source text

In this section we describe our main contribution, Adversarial Optimism (AdOpt) in detail. At a high level, AdOpt uses two classifiers. The first one is a classifier trained on all the accepted data thus far, without any de-biasing. We refer to this as the “biased” classifier hereafter. The second one is our adversarially de-biased classifier trained on the same data. AdOpt then proceeds as follows:

An example of annotated data

Type: Concision

Make this paragraph shorter.

Revised text

In this section we describe our main contribution, Adversarial Optimism (AdOpt) in detail. AdOpt uses two classifiers. The first one is a “biased” classifier trained on all the accepted data thus far. The second one is our adversarially de-biased classifier. AdOpt then proceeds as follows:

Statistics

- Rewriting light
- Rewriting medium
- Rewriting heavy
- Development
- Content addition
- Content substitution
- Concison
- Content deletion
- Unusable

10 annotators

47 796 paragraphs in total

641 annotated paragraphs (73.32% with cross annotation)

Agreement: 0.499 (strict), 0.693(merged) with Krippendorff's alpha

# instructions	0	1	2
# paragraphs	327	56	258

Perspectives

- Automatically annotate the train data
- Fine tune a specialised small open source model
- Compare it with foundation LLM (performances and energy)

Experiments

Do the instructions improve the revision?

Metric	rougeL		sari		bert-score	
	Label	Inst	Label	Inst	Label	Inst
no edits	78.49		60.69		95.98	
coedit-xl	67.50	67.70	39.07	39.20	93.51	93.59
Mistral-7B-Instruct-v0.2	46.26	48.57†	28.64	30.55†	91.45	91.81†
Llama-3-8B-Instruct	50.78	55.96†	30.70	35.15†	91.89	92.71†
GPT4o	57.99	66.17†	33.33	41.39†	92.89	94.11†