



**HAL**  
open science

## Active learning to measure opinion and violence in French newspapers

Paul Guélorget, Guillaume Gadek, Titus Zaharia, Bruno Grilheres

► **To cite this version:**

Paul Guélorget, Guillaume Gadek, Titus Zaharia, Bruno Grilheres. Active learning to measure opinion and violence in French newspapers. *Procedia Computer Science*, 2021, 192, pp.202-211. 10.1016/j.procs.2021.08.021 . hal-03595547

**HAL Id: hal-03595547**

**<https://hal.science/hal-03595547v1>**

Submitted on 16 Oct 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



25th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems

## Active learning to measure opinion and violence in French newspapers

Paul Guélorget<sup>a,b</sup>, Guillaume Gadek<sup>a</sup>, Titus Zaharia<sup>b</sup>, Bruno Grilheres<sup>a</sup>

<sup>a</sup>*Airbus Defence and Space, 1 Boulevard Jean Moulin, 78990 Élanecourt, France*

<sup>b</sup>*Institut Polytechnique de Paris, Institut Mines-Télécom, Télécom SudParis, 9 Rue Charles Fourier, 91000 Évry, France*

### Abstract

News articles analysis may be oversimplified when restricted to detecting classes of interest already benefiting from trustworthy labeled datasets, like political affiliation or fakeness. Behind an apparent neutrality, an editorial slant may be embodied by favoring one-sided interviews, avoiding topics or choosing oriented illustrations. These challenges, seen as machine learning problems, would require a tedious annotation task. We introduce ReALMS, an active learning framework capable of quickly elaborating models which detect arbitrary classes in multi-modal text and image documents. Evidence of this capability is given by a case study on French news outlets: the detection of subjectivity, demonstrations and violence.

© 2021 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Peer-review under responsibility of the scientific committee of the KES International.

*Keywords:* active learning; text and image classification; multimodal classification; media analysis

### 1. Introduction

The last five years have seen the worldwide dissemination of the *fake news* buzzword, associated with various phenomena in politics and in the media. The role of journalism as information providers also includes a part of *attention retainer*, with some news outlets apparently renouncing to information (replacing it with *contents* or *stories*). The machine learning domain already proposes great algorithms to tag news articles as “normal” or as “fake-looking”. Limiting the analysis to single documents, classic approaches are automatic fact verification [2] and classification based on stylistic features [13, 8]. While successfully detecting garbage news, these works seem to dismiss the fact that news outlets are entitled to promote opinions, tribunes, open letters and editorials.

Political science has a long history of studying the political leaning of news outlets [23], counting the number of articles per topic and scrutinizing the coverage of events with regard to their impact in an election. These analyses are

\* Paul Guélorget

*E-mail address:* [paul.guelorget@telecom-sudparis.fr](mailto:paul.guelorget@telecom-sudparis.fr)

particularly focused on the United States bipartisan political system, where supporters of a party must be opponents of the other. However, this does not reflect the realities of many other countries, cultures and languages.

We have the intuition that the ability to detect and study arbitrary classes could contribute to a better understanding of the media landscape. In this article, we propose to consider three new tasks of French press analysis: the classification of articles relaying one-sided opinions and the detection of two classes, demonstration and violence, in title/image article previews. We instantiate these classifiers through a novel active learning framework, named ReALMS, agnostic to the labeling user’s availability, thus bringing substantial advantages in terms of the effective time spent to obtain a satisfying classification model.

This paper is structured as follows: Section 2 gives some details about the news article classification problems and describes previous approaches. Then we introduce ReALMS (Real-time Active Learning as Micro-Services), our active learning framework in Section 3. Section 4 presents the collection of Reddit data used to carry out our real life case study of French press articles and the training of ReALMS on this data. The obtained models are put to contribution to characterize news outlets in a novel manner. Finally, Section 5 concludes this paper.

## 2. Related work: news articles and active learning

### 2.1. Analyzing media and politics

An intuitive scale for media analysis consists in measuring the political leaning of all major US news outlets along a “liberal-conservative” axis, which can be based on newspaper headlines, using sentiment analysis, and taking into account economic events [19]. This kind of investigation has also been performed while limited to *editos*, that is excluding objective, purely informative articles [11]. Media bias can be measured without reading any article: a way to do so is to rely on demographics data supplied by advertisement companies such as Facebook [24], resulting in an efficient qualification of news outlets along ideological (liberal-conservative), identity (age, race, gender) and income. Another way is to follow the *retweet* links on a social network, tagging each news outlet account with a red/blue color [30]. The political non-binary stance attribution has only been the focus of a few work; only recently have some datasets been constituted, opening this as a frame of social network analysis [7].

Previous machine learning based analysis provided invaluable insight. First, that most news outlets are frequently “informative only” (i.e., non partisan). Second, that the agenda setting behavior is probably more subtle than initially thought as “*little evidence exists of systematic differences in story selection*” [4]. The growing success of AI techniques results in attempts to rationalize the analysis of newspapers in a current context of “fake news”: either to detect fake pieces of information (fact-checking) [2], or stylistic features correlated with low-quality journalism [8]. The importance of stylistic features and emotion intensity has been underlined in a recent survey work [31], suggesting that “persuasive” news outlets prefer to insist on impacting content, up to fear-inducing, such as raw violence.

Violence detection is often seen as a video analysis task, with a direct exploitation concerning movies [14]. For textual news articles, some resources exist, mostly focusing on the topic of gun violence in the United States; they enable topic classification and entity identification [21]. Another stream of work investigates the “framing”, i.e. the main aspect of gun violence that is relayed by a given article, instantiated as a classification task (Politics, Economic consequences, 2nd Amendment, Gun control, Ethnicity, ...) [18].

### 2.2. The need for active learning

Automating this analysis is possible as current technologies propose effective methods to process text and image modalities. The processing of textual data is usually addressed by two main processes: feature extraction and then classification. Widespread text feature extraction methods are pre-computed word embeddings according to which words are represented by fixed-length vectors such as FastText[12], ELMo [22] and BERT [6]). Classification usually relies on Recurrent and/or Convolutional Neural Networks (RNN, CNN) architectures. Images are also present in news articles and are a significant part of their embedded previews on social media. A very popular solution to image classification is transfer learning [20]. CNN architectures such as ResNet [10] or EfficientNet [26] offer performing and versatile feature extractor trunks when pre-trained with very broad datasets like ImageNet [25] beforehand.

The aforementioned text and image classification methods often rely on a significant number of labeled instances, ranging from thousands to millions. Evolving trends and the diversity of taxonomies make so that many problems lack or will lack the reliable datasets that could ease their solving. A solution consists in considering the active learning process [27, 1]. During an active learning iteration, new documents are either synthesized or drawn from a pool – or stream – of unlabeled documents according to criteria dictated by a sampling strategy. They are then presented to a human expert – or oracle – for labeling. The key principle of active learning consists in selecting the examples which are the most likely to benefit training. Multiple sampling criteria can be considered, such as uncertainty [15], information density (or representativeness) [16], diversity [3] or query-by-committee [5].

### 2.3. Discussion

First, we estimate that too few work explore the non-US, non-bipartisan setups. Most systems are specifically designed for English language news analysis, while most of the world does rely on non-English “local” media. Thus we propose to explore the French media landscape through one online social entry point, the Reddit r/france page.

Second, most of the related work requires comprehensive expert-annotated datasets, which thoroughly limits the possible questions. To empower the analyst to feel free to ask and research whatever question they need, we introduce ReALMS, an active learning framework, enabling to sketch and refine analytics from non-annotated corpora and datasets. The use cases displayed in this work showcase the opportunities offered by active learning in general and by our proposed framework in particular.

## 3. Real-time active learning as micro-services

Usually, the evaluation of an active learning method is based on a separate testing pool: by measuring a performance metric (e. g. accuracy) after consuming a labeling budget, or by counting the required labeling steps before a desired metric threshold is reached. More generally, active learning performance can be assessed by the concavity of the learning curve, which plots the performance metric as a function of the (growing) training set. Granted that the sole number of labeling is a determining cost, it alone conceals crucial drawbacks of active learning when conducted in real life conditions: (a) counting with learning cycles ignores the actual time spent by the expert; (b) in real-life experimental conditions, there exists no independent validation set that can be used to measure the improvement of the trained model; (c) hyper-parameters setting has to be valid once and for all.

To address these drawbacks, in this paper, we propose ReALMS (Real-time Active Learning as Micro-Services), a technical solution in the form of a set of cooperating micro-services, carrying out the tasks required by active learning: receive documents and labels, train neural networks, update document descriptions, and sample documents that are candidates for labeling. Our proposed solution includes a variety of supervised models applied to a variety of modalities: as part of this work, on text and on multi-modal (text and image) classification. Furthermore, we claim that our proposed technical solution can be quickly integrated in a broader media analysis / business intelligence framework.

### 3.1. Notations

To solve multiclass or multi-label classification problem with  $c$  classes, in an active learning setting,  $\mathcal{U} = \{(x_i, \hat{x}_i, \hat{y}_i)\}_{i \leq u}$  is the set of all unlabeled documents stored in the framework, represented by their frozen embeddings  $x$ , their deep vector representation  $\hat{x}$  and prediction scores  $\hat{y} \in \mathbb{R}^c$  (cf. section 3.2.2.)  $\mathcal{L} = \{(x_i, \hat{x}_i, y_i)\}$  is the set of all labeled documents stored in the framework, where  $y \in \{0, 1\}^c$  is the ground-truth label provided by an expert oracle.  $\mathcal{L}$  is split into training set  $\mathcal{T}$  and validation set  $\mathcal{V}$ . The best classification model at a given moment is noted  $\mathcal{M}$ , while a candidate model being trained is noted  $M$ . For the sake of simplicity, we liken a document to its frozen embedding  $x$ .

### 3.2. Proposed organization

Our proposed framework aims to parallelize time consuming tasks that are usually conducted sequentially: training, sampling and description of unlabeled documents. Moreover, we blur the boundary between pool-based (sampling

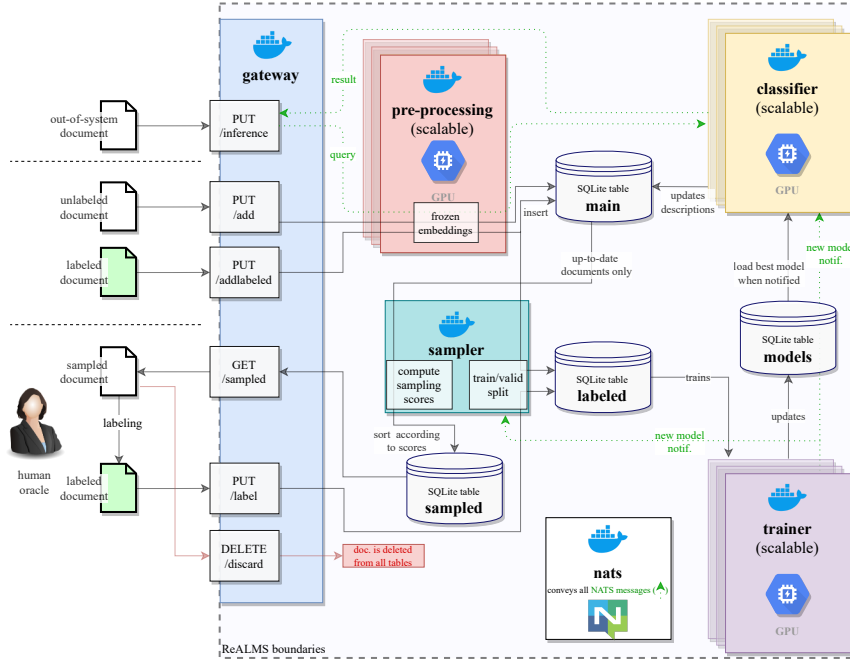


Fig. 1. ReALMS organization. Several docker services cooperate to carry out all time-consuming tasks required by active learning.

from a static pool of unlabeled documents) and stream-based (sample or discard incoming documents) active learning scenario by allowing the continuous reception of new documents during the lifespan of the system. The overall organization of our framework is illustrated on Figure 1. The following sections describe the networked micro-services.

### 3.2.1. Pre-processing service.

The pre-processing service is in charge of computing frozen vector embeddings  $x$  for all incoming raw documents. Unless deleted, a document embedding is stored during the whole lifespan of the system in the *main* table of the database (displayed on Figure 1) to avoid multiple re-computing. The trainer service uses the frozen embeddings  $\{x_i\}$  as inputs once per training epoch to train a neural network (cf. trainer service); the classifier service uses the frozen embeddings  $\{x_i\}$  to update the representation of documents (cf. classifier service). Text inputs are transformed into a list of FastText 300-dimensional embeddings of their tokens. Input images are resized to  $380 \times 380$  pixels and fed to a pre-trained EfficientNet-B4 trunk producing 1792-dimensional embeddings. This service is scalable: in case of a heavy stream of incoming documents, several instances can subscribe to the same NATS queue.

### 3.2.2. Classifier service.

In active learning, the classification model  $\mathcal{M}$  is intended to be upgraded several times. The classifier service always has to keep up-to-date the description of all documents: prediction results  $\{\hat{y}_i\}$  and deep feature vectors  $\{\hat{x}_i\}$ . It is implemented as a Tensorflow container with GPU support. The service accesses the *main* table representing the pool of registered documents, represented by their frozen embeddings  $\{x_i\}$ . It is notified whenever a new model version is discovered by the trainer service. When this occurs, the parameters of the new model  $\mathcal{M}$  are loaded and the service updates the documents whose descriptions are obsolete, by random batches. Also, the classifier service answers on-demand inference queries from the user. Several instances can run simultaneously in separate containers to speed up the description update. From now on, documents whose description is up-to-date are called *up-to-date documents*; others are called *obsolete documents*.

### 3.2.3. Sampler service.

**Sampling task.** The sampler service relies on the up-to-date document descriptions  $\{(\hat{x}_i, \hat{y}_i)\}$  produced by the classifier service to compute sampling scores. The highest sampling score document will be the next to be labeled by the

Table 1. Sampling criteria.

Criteria	Definition	Formula
Uncertainty	the prediction vector’s entropy	$c_u(x_i) = -\sum_{i=1}^c \hat{y}_i \log \hat{y}_i$
Representativeness	opposite of the mean distance of a document to its nearest neighbors	$c_r(x_i) = 1 - \frac{1}{n} \sum_{j=1}^n \text{dist}(\hat{x}_i, \hat{x}_j)$
Diversity	smallest distance between an unlabeled document and the already labeled documents	$c_d(x_i) = \min_{x_j \in \mathcal{L}} \text{dist}(\hat{x}_i, \hat{x}_j)$
Global	product of sub-criteria $c_u$ , $c_r$ and $c_d$	$C(x_i) = c_u(x_i) \times c_r(x_i) \times c_d(x_i)$

oracle. Only up-to-date documents are candidates for sampling. Unless the list of up-to-date documents has ceased to expand, the sampler continuously carries out the sampling task. This mechanism ensures that the incompleteness of the description task is not an obstacle to sampling.

The sampling criterion according to which unlabeled documents are sorted is the combination of three sub-criteria: uncertainty  $c_u$ , representativeness  $c_r$  and diversity  $c_d$ , which are defined in Table 1. The distance used is cosine distance. Up-to-date unlabeled documents are sorted according to their global sampling scores, decreasingly, to populate a sampling queue.

*Training/validation split strategy.* The split decision is performed when a document is labeled, and it is final: a document marked as *validation* will be used for validation by all future models. ReALMS is initialized with a target validation ratio  $r_V$  dictating the proportion of labeled documents to use as a validation set  $\mathcal{V}$ , isolated from the training set  $\mathcal{T}$ . When a labeled document is received, its training/validation split is determined by its label and the current state of the validation set. The incoming document is assigned to  $\mathcal{V}$  if  $\mathcal{V}$  is lacking members of the received document’s classes  $y = \{y_1, \dots, y_c\} \in \{0, 1\}^c$ .

$$c_V(y) = \min_{i \in \llbracket 1, c \rrbracket} \frac{|\{y' \in \mathcal{V}; y'_i = y_i\}|}{|\{y' \in \mathcal{V} \cup \mathcal{T}; y'_i = y_i\}|} \quad (1)$$

Thus defined,  $c_V$  accounts for the validation ratio of the most depleted label (positive or negative) in  $\mathcal{V}$ . If  $c_V(y) < r_V$ , the received document is assigned to validation.

#### 3.2.4. Trainer service.

*Training routine.* The trainer service exploits  $\mathcal{T}$  to produce a model that performs best on  $\mathcal{V}$ . Each trained model is compared to *best* model  $\mathcal{M}$ . Because the validation set  $\mathcal{V}$  grows over time, the f1-score obtained by best model  $\mathcal{M}$  at the time of its creation is obsolete. For this reason, we re-evaluate  $\mathcal{M}$  on the current state of  $\mathcal{V}$ . Several trainer containers can run simultaneously to speed up the discovering of new best models. ReALMS is initialized with a random architecture ratio  $r_R \in [0, 1]$  dictating the probability of starting a new training from scratch with a new model  $M$  or to re-load and resume the training of  $\mathcal{M}$ . Also, confident documents may be added to  $\mathcal{T}$  for training [29].

*Model architectures.* Figure 2 represents the overall architecture of the models used in our framework. For mono-modal classification, the fusion step is bypassed. Layers outputs are regularized by dropout and weight decay and are ReLU-activated, with the exception of the final classification layer using a softmax (respectively sigmoid) activation for multiclass (respectively multi-label) classification. We use a random number (between 0 and 4) of 1-d convolutional layers to process the text modality, each with a random number of filters (between  $2^6$  and  $2^{10}$ ) and a random kernel size (3 or 5). The global pooling method is chosen randomly between global max pooling, global average pooling and self-attention [17]. Text and image feature vectors are processed by a random number of dense layers (between 0 and 4) with a random number of units (between  $2^6$  and  $2^{10}$ ), whose output are merged following one randomly chosen fusion method out of late fusion, most certain and a multi-layer CentralNet[28] sub-network.

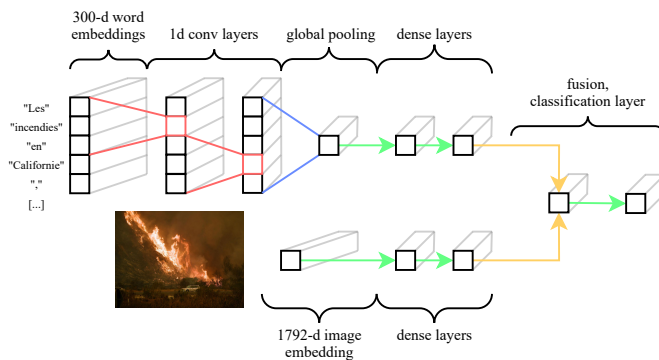


Fig. 2. Template of models used in ReALMS for multi-modal text/image classification

### 3.2.5. Helper services.

As shown on Figure 1, ReALMS relies on additional services. The gateway service exposes a web interface to manage all incoming documents and labels, enabling a continuous reception. It also exposes a labeling GUI as a web page. Nats service is a container running message broker NATS<sup>1</sup>, conveniently conveying asynchronous queries and notifications (as dotted green arrows on Figure 1).

## 4. Application on French politics

French politics present a few differences with its US counterparts, in terms of language (French) and political landscape structure. Stance is not a binary classification task, because the center-right government is criticized from both sides by left and right wings oppositions. However, as in the USA, most news articles are informative although their political leaning is perceivable from the addressed topics. In this article, we propose to compare the emission and engagement received by objective versus subjective articles. We believe this may help to qualify the behavior of news outlets.

### 4.1. Data collection

We aggregated data from the *r/france subreddit*<sup>2</sup> by targeting submissions that embed URLs from a list of 25 French news outlets, using the Pushshift API<sup>3</sup>. The titles, bodies and main images of the embedded articles are retrieved using NewsPlease [9]. The number of top-level comments per submission is retrieved with the Python Reddit Application Programming Interface (API) Wrapper<sup>4</sup>. Once duplicates have been removed, we have a total of 39,611 news articles. The resulting corpora are published<sup>5</sup> in order to facilitate future research on these topics.

### 4.2. Detection of articles conveying one-sided opinions

This text classification problem consists in detecting articles that relay one-sided opinions, such as interviews, one-sided paraphrased opinions and straight tribunes. For this first use case, we chose to restrict our study to news article marked with the “Politics” flair. Input documents consist in the junction of their titles and bodies. This problem is tackled as a binary classification task. Target validation ratio  $r_V$  is set to 1/5. During labeling, articles dealing with

<sup>1</sup> NATS, <https://docs.nats.io/>

<sup>2</sup> [reddit.com/r/france](https://reddit.com/r/france)

<sup>3</sup> [github.com/pushshift/api](https://github.com/pushshift/api)

<sup>4</sup> PRAW: The Python Reddit Api Wrapper, <https://praw.readthedocs.io/en/v3.0.0/>

<sup>5</sup> <https://conf.telecom-sudparis.eu/realms/>

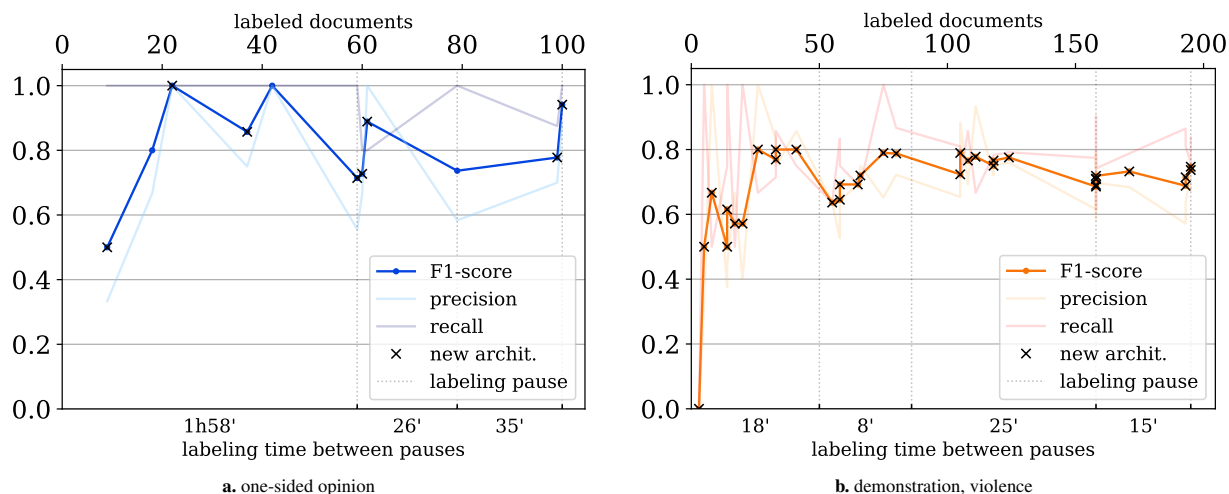


Fig. 3. Learning curves for one-sided opinion detection (left) and demonstration/violence detection (right). “New archit.” signifies that the model was obtained with a new, random architecture, as opposed to improving the already existing model.

foreign politics are discarded. Figure 3.a shows the learning curves obtained while sampled articles were manually labeled by the authors. Vertical dotted lines stand for hours-long interruptions of the labeling task. After labeling 100 articles, our model reached 0.941 f1-score (0.889 precision, 1.0 recall) on 20 validation articles. The final neural network’s architecture is  $\{conv1d(137 \text{ filters of size } 3), conv1d(124 \text{ filters of size } 3), global \text{ average pooling}, dense(66 \text{ units})\}$ . Table 2 contains titles of unlabeled articles whose prediction scores are the strongest. The obtained model is able to retrieve articles demonstrating pure subjective content yet published in the form of press articles. Testing on 20 independent documents yields a 0.82 f1-score.

Table 2. Most positive and negative results for label “one-sided opinion” (from unlabeled corpus)

prediction	translated title
0.990	François Hollande warns against Donald Trump’s victory that many think is impossible
0.986	Worried about Mélenchon’s breakthrough, Hollande calls for “renewal”
0.985	VIDEO. Alain Juppé “screws” those who find him “very conventional”
0.008	Presidential election: Macron at 24.01%, Le Pen at 21.30%, according to definitive first round results
0.008	Presidential election: 9.8 million French people followed the debate on TF1
0.005	Brexit: the (simple) graphic to summarize five months of procrastination between London and Brussels

#### 4.3. Demonstration and violence

Our second use case consists in classifying text/image pairs belonging to zero, one or two of arbitrary “demonstration” and “violence” categories. Articles lacking a valid image are discarded. We used keywords to quickly filter on a controversial topic about police violence: articles are retained if their body contains one or more of the following keywords: *manifestation*, *manifs*, *sécurité globale* (a controversial French bill), *violence(s) policière(s)* (police violence), *CRS* (French riot police), *émeute* (riot), *dégradations* (degradation). We use title/image pairs as they are the constituents of the previews displayed on social media.

This text and image classification problem consists in detecting articles previews that mention or depict demonstration and/or violence. This problem is tackled as a multi-label classification task. Target validation ratio  $r_V$  is set to 1/4. During labeling, articles dealing with foreign politics are discarded. Figure 3.b shows the learning curves obtained while sampled articles were manually labeled by the authors. Vertical dotted lines stand for hours-long interruptions of the labeling task.

After labeling 195 articles, our model reached 0.747 micro f1-score (0.674 micro precision, 0.838 micro recall) on 50 validation articles. The final architecture performs a late fusion between the text branch  $\{conv1d(89 \text{ filters of size } 5), conv1d(635 \text{ filters of size } 3), global \text{ average pooling}, dense(140), dense(72), dense(122)\}$  and the image branch



Table 3. Influence of input modalities on prediction scores









translated input text	input image	
		
Unemployment: the number of job seekers fell sharply in 2019	demonstration: 0.091 violence: 0.086	demonstration: 0.473 violence: 0.504
Evacuation of the Sully bridge: the CRS commander himself lost consciousness "by suffocation of tear gas"	demonstration: 0.561 violence: 0.577	demonstration: 0.943 violence: 0.994

Table 4. Highest predictions scores for label “demonstration” (left) and “violence” (right), taken from unlabeled corpus

image	translated title	image	translated title
	EXCLUSIVE. <i>Black bloc, féminicide, dégaisme...</i> The new words of the Larousse		A same-sex couple assaulted by about ten armed young people in Seine-Saint-Denis
	Yellow jackets, act 9: towards an action in “the center of France” or in La Défense?		Dallas: five police officers killed, suspects arrested, another is dead
	LIVE. Pension reform: unions call for local actions on Thursday 12 December and for a national demonstration on Tuesday 17 December		“Purge” against police officers: the caller placed in custody

{dense(65), dense(367), dense(75), dense(81)}. Results in Table 3 confirm that both input modalities have an impact on prediction results. Table 4 contains titles and images of unlabeled articles previews whose prediction scores are the strongest. The model can still be misled by words such as “black bloc” (the name of a demonstration tactic). Testing on 20 independent documents yields a 0.70 f1-score.

#### 4.4. Exploitation of the predictive capability for media analysis

We perform a case study in two steps: first as a text classification task, detecting one-sided opinion articles, quantifying their impact and qualifying their sources. Second, as a multi-modal identification task limited to a political subtopic about public demonstrations: we propose to qualify the news outlets by their display of violence to qualify these events.

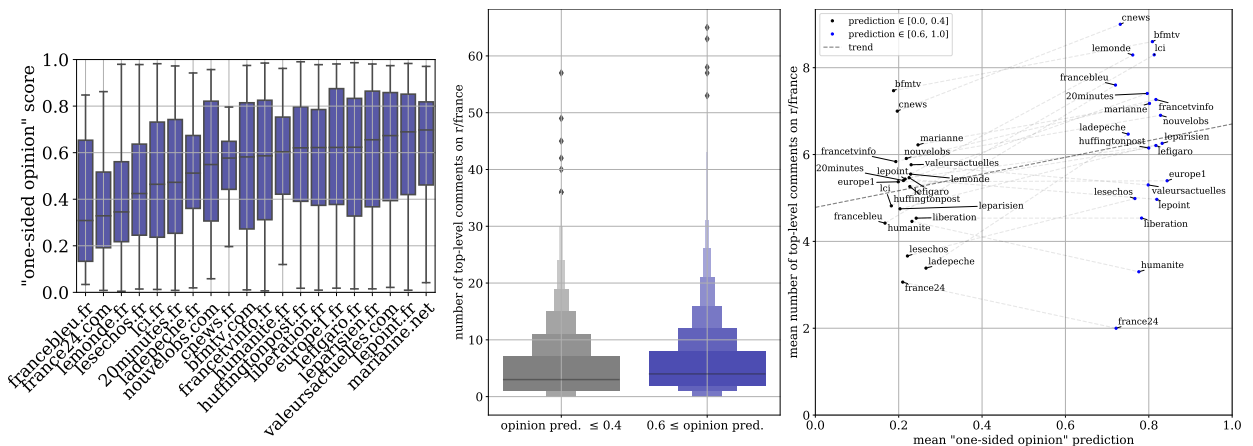


Fig. 4. Opinion score distribution across news outlets (left), mean number of comments between factual and opinion-relaying articles (center) and per news outlet (right)

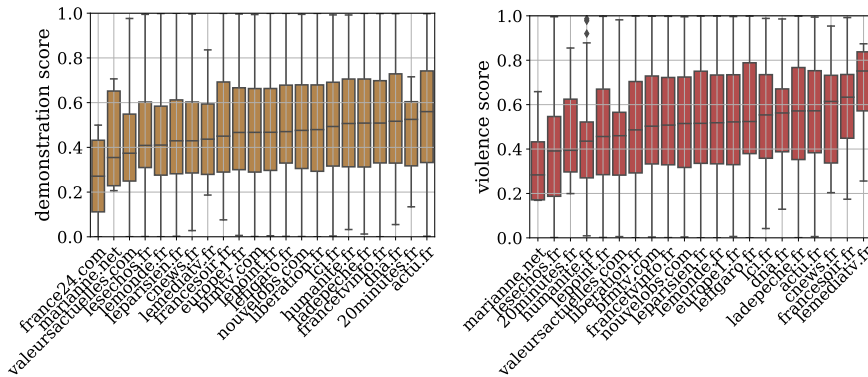


Fig. 5. Demonstration score distribution of article previews (left) and violence score distribution of 8 article previews detected as belonging to the *demonstration* class (right), across news outlets

One of the working questions bears on the fact that newspapers produce both *informative* or *opinion-sided* articles. Each article receives a score of opinion, ranging from 0 (informative) to 1 (expressing one person’s opinion). On Figure 4, these scores are aggregated by news outlet, and sorted by their median score. *FranceBleu*, the local public information channel, can be qualified as informative while *Le Point* and *Marianne*, at the other side of the chart, are better diffused on Reddit when they relay opinions.

Subjective articles are often supposed to be quicker to spread on social networks. Figure 4 answers this question: we arbitrarily split the documents as “factual” (opinion score below 0.4, in grey) and “opinion-relaying” (score over 0.6, in blue). The global trend is indeed for opinion-relaying articles to receive some more top-level comments; this trend is however not very strong. Another phenomenon to underline is the great diversity between news outlets: in terms of engagement, the continuous information TV channels *LCI* and *BFM* seem to consistently trigger more engagement. At the extreme opposite, *france24* does not receive as many comments; this channel is also more watched abroad than inside France. Among the news outlets for which there is no gain in engagement when their articles are relays of subjectivity, *Valeurs actuelles*, *Le Point* and *L’Humanité* have a reputation of being relays of their political factions. It seems that they do not gain more attention when explicitly relaying an opinion.

At the end of 2020, a security law proposal in France triggered demonstrations, including violent scenes. News outlets chose different editorial paths, arguably to encourage or dissuade people to gather in the streets. Figure 5 displays the news outlets, sorted according to their median scores (whether the articles covers demonstrations). We propose a similar view about violence, for the articles that are predicted as covering demonstrations, also in Figure 5. Indeed, a well spread narrative suggests that all demonstrations always result in violence and riots whether it would be due to the demonstrators or to repression: this chart brings some quantifiable insights on this topic. Interestingly, the left-wing outlet *Le Média* known for publishing field reports of demonstrations has the highest median violence score, while formerly left-leaning outlet *Marianne* is the least prone to mention violence in its articles tackling demonstrations.

## 5. Conclusion and future work

In this paper, we introduced an efficient implementation for Active Learning, and demonstrate its relevance on two classification tasks in a setup of political media analysis. The framework is highly scalable and solves many pain points, notably the parallelization of time-consuming tasks. The case study on French language articles posted on Reddit illustrates the possibilities given by this technology: it facilitates the development of AI models not only for English, and not limited to one modality. We believe that active learning will help practitioners to adopt AI in their analysis, enabling them to explore more research questions and to quantify more subtle phenomena that do not already benefit from a large and reliable labeled dataset. Improvement leads are twofold: first in the AI domain, to take into account the recent progress in explainability, ensuring that the models learns with the correct rationales. Second, in the media analysis domain, we would like to explore and to deepen our analysis, in order to continue qualifying and characterizing the main information sources of the citizen.

## References

- [1] Angluin, Dana, 1988. Queries and Concept Learning. *Machine Learning* 2, 319–342.
- [2] Atanasova, Pepa, Nakov, Preslav, Márquez, Luís, Barrón-Cedeño, Alberto, Karadzhov, Georgi, Mihaylova, Tsvetomila, Mohtarami, Mitra, and Glass, James, 2019. Automatic fact-checking using context and discourse information. *Journal of Data and Information Quality (JDIQ)* 11, 1–27.
- [3] Brinker, Klaus, 2003. Incorporating Diversity in Active Learning with Support Vector Machines, in: *Proceedings of the Twentieth International Conference on International Conference on Machine Learning*, AAAI Press, Washington, DC, USA. pp. 59–66.
- [4] Budak, Ceren, Goel, Sharad, and Rao, Justin M, 2016. Fair and balanced? quantifying media bias through crowdsourced content analysis. *Public Opinion Quarterly* 80, 250–271.
- [5] Dagan, Ido, and Engelson, Sean P., 1995. Committee-Based Sampling For Training Probabilistic Classifiers, in: *In Proceedings of the Twelfth International Conference on Machine Learning*, Morgan Kaufmann. pp. 150–157.
- [6] Devlin, Jacob, Chang, Ming-Wei, Lee, Kenton, and Toutanova, Kristina, 2018. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv:1810.04805 [cs] ArXiv: 1810.04805.
- [7] Fraiser, Ophélie, Cabanac, Guillaume, Pitarch, Yoann, Besancon, Romaric, and Boughanem, Mohand, 2018. # élysée2017fr: The 2017 french presidential campaign on twitter, in: *Proceedings of the International AAAI Conference on Web and Social Media*.
- [8] Gadek, Guillaume, and Guélorget, Paul, 2020. An interpretable model to measure fakeness and emotion in news. *Procedia Computer Science* 176, 78–87.
- [9] Hamborg, Felix, Meuschke, Norman, Breitingner, Corinna, and Gipp, Bela, 2017. news-please: A Generic News Crawler and Extractor, in: *Proceedings of the 15th International Symposium of Information Science*, pp. 218–223. Event-place: Berlin.
- [10] He, K., Zhang, X., Ren, S., and Sun, J., 2016. Deep Residual Learning for Image Recognition, in: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778.
- [11] Ho, Daniel E, Quinn, Kevin M, et al., 2008. Measuring explicit political positions of media. *Quarterly Journal of Political Science* 3, 353–377.
- [12] Joulin, Armand, Grave, Edouard, Bojanowski, Piotr, and Mikolov, Tomas, 2016. Bag of Tricks for Efficient Text Classification. arXiv:1607.01759 [cs] ArXiv: 1607.01759.
- [13] Katsaros, Dimitrios, Stavropoulos, George, and Papakostas, Dimitrios, 2019. Which machine learning paradigm for fake news detection?, in: *2019 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, IEEE. pp. 383–387.
- [14] Khan, Samee Ullah, Haq, Ijaz Ul, Rho, Seungmin, Baik, Sung Wook, and Lee, Mi Young, 2019. Cover the violence: A novel deep-learning-based approach towards violence-detection in movies. *Applied Sciences* 9, 4963.
- [15] Lewis, David D., and Catlett, Jason, 1994. Heterogeneous Uncertainty Sampling for Supervised Learning, in: *In Proceedings of the Eleventh International Conference on Machine Learning*, Morgan Kaufmann. pp. 148–156.
- [16] Li, X., and Guo, Y., 2013. Adaptive Active Learning for Image Classification, in: *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 859–866.
- [17] Lin, Zhouhan, Feng, Minwei, Santos, Cicero Nogueira dos, Yu, Mo, Xiang, Bing, Zhou, Bowen, and Bengio, Yoshua, 2017. A Structured Self-attentive Sentence Embedding. arXiv:1703.03130 [cs] ArXiv: 1703.03130.
- [18] Liu, Siyi, Guo, Lei, Mays, Kate, Betke, Margrit, and Wijaya, Derry Tanti, 2019. Detecting frames in news headlines and its application to analyzing news framing trends surrounding us gun violence, in: *Proceedings of the 23rd CoNLL Conference*, pp. 504–514.
- [19] Lott, John R, and Hassett, Kevin A, 2014. Is newspaper coverage of economic events politically biased? *Public choice* 160, 65–108.
- [20] Pan, Sinno Jialin, and Yang, Qiang, 2010. A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering* 22, 1345–1359.
- [21] Pavlick, Ellie, Ji, Heng, Pan, Xiaoman, and Callison-Burch, Chris, 2016. The gun violence database: A new task and data set for nlp, in: *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pp. 1018–1024.
- [22] Peters, Matthew E., Neumann, Mark, Iyyer, Mohit, Gardner, Matt, Clark, Christopher, Lee, Kenton, and Zettlemoyer, Luke, 2018. Deep contextualized word representations. arXiv:1802.05365 [cs] ArXiv: 1802.05365.
- [23] Puglisi, Riccardo, 2011. Being the new york times: the political behaviour of a newspaper. *The BE journal of economic analysis & policy* 11.
- [24] Ribeiro, Filipe, Henrique, Lucas, Benevenuto, Fabricio, Chakraborty, Abhijnan, Kulshrestha, Juhi, Babaei, Mahmoudreza, and Gummadi, Krishna, 2018. Media bias monitor: Quantifying biases of social media news outlets at large-scale, in: *Proceedings of the International AAAI Conference on Web and Social Media*.
- [25] Russakovsky, Olga, Deng, Jia, Su, Hao, Krause, Jonathan, Satheesh, Sanjeev, Ma, Sean, Huang, Zhiheng, Karpathy, Andrej, Khosla, Aditya, Bernstein, Michael, Berg, Alexander C., and Fei-Fei, Li, 2014. ImageNet Large Scale Visual Recognition Challenge. arXiv:1409.0575 [cs] .
- [26] Tan, Mingxing, and Le, Quoc, 2019. Efficientnet: Rethinking model scaling for convolutional neural networks, in: *International Conference on Machine Learning*, PMLR. pp. 6105–6114.
- [27] Valiant, L. G., 1984. A Theory of the Learnable, in: *Proceedings of the Sixteenth Annual ACM Symposium on Theory of Computing*, ACM, New York, NY, USA. pp. 436–445.
- [28] Vielzeuf, Valentin, Lechervy, Alexis, Pateux, Stéphane, and Jurie, Frédéric, 2018. CentralNet: a Multilayer Approach for Multimodal Fusion. arXiv:1808.07275 [cs] ArXiv: 1808.07275.
- [29] Wang, Keze, Zhang, Dongyu, Li, Ya, Zhang, Ruimao, and Lin, Liang, 2016. Cost-Effective Active Learning for Deep Image Classification. *IEEE Transactions on Circuits and Systems for Video Technology* 27, 1–1.
- [30] Wong, Felix Ming Fai, Tan, Chee Wei, Sen, Soumya, and Chiang, Mung, 2013. Quantifying political leaning from tweets and retweets, in: *Proceedings of the International AAAI Conference on Web and Social Media*.
- [31] Zhou, Xinyi, and Zafarani, Reza, 2018. Fake news: A survey of research, detection methods, and opportunities. arXiv:1812.00315 .