Group 13

Balancing Political Bias in News Article Retrieval

Thomas Jahoda
- IR Pipeline & evaluation
- Political balance metric & ranking correction methods
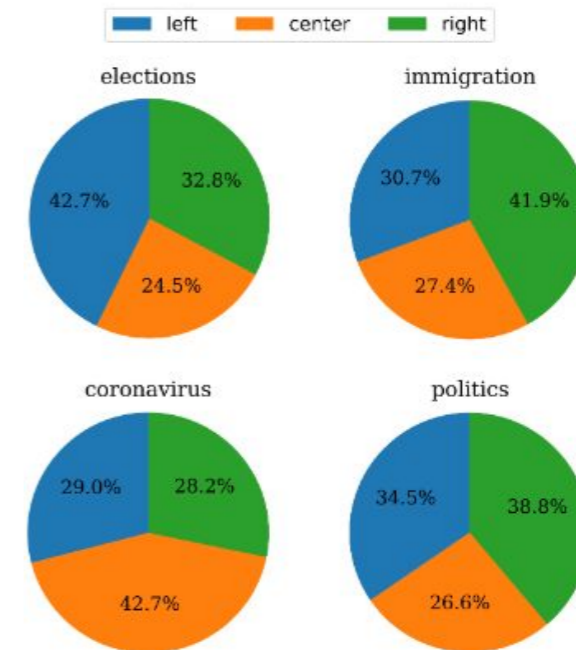
Raphael Rogi
- Dataset analysis & preprocessing
- Bias-classifier

Project Repository

Motivation

- The viewpoint of authors and news outlets influences how a story is told
- Consuming news only from one side can lead to extreme political ideology
- It is important to make the reader aware of political bias
- Balance bias in search results

Ramy Baly et al., We Can Detect Your Bias: Predicting the Political Ideology of News Articles
Israel headlines from https://www.allsides.com [06.01.2024]

Datasets

all-the-news dataset 🤗

year	month	day	author	title	article	url	section	publication
2016	12.0	9	Lee Drutman	We should take concerns about the health of li...	This post is part of Polyarchy, an independent...	https://www.vox.com/polyarchy/2016/12/9/138983...	None	Vox
2016	10.0	7	Scott Davis	Colts GM Ryan Grigson says Andrew Luck's contr...	The Indianapolis Colts made Andrew Luck the h...	https://www.businessinsider.com/colts-gm-ryan-...	None	Business Insider
2018	1.0	26	None	Trump denies report he ordered Mueller fired	DAVOS, Switzerland (Reuters) - U.S. President ...	https://www.reuters.com/article/us-davos-meeti...	Davos	Reuters
2019	6.0	27	None	France's Sarkozy reveals his 'Passions' but in...	PARIS (Reuters) - Former French president Nico...	https://www.reuters.com/article/france-politic...	World News	Reuters
2016	1.0	27	None	Paris Hilton: Woman In Black For Uncle Monty's...	Paris Hilton arrived at LAX Wednesday dressed ...	https://www.tmz.com/2016/01/27/paris-hilton-mo...	None	TMZ

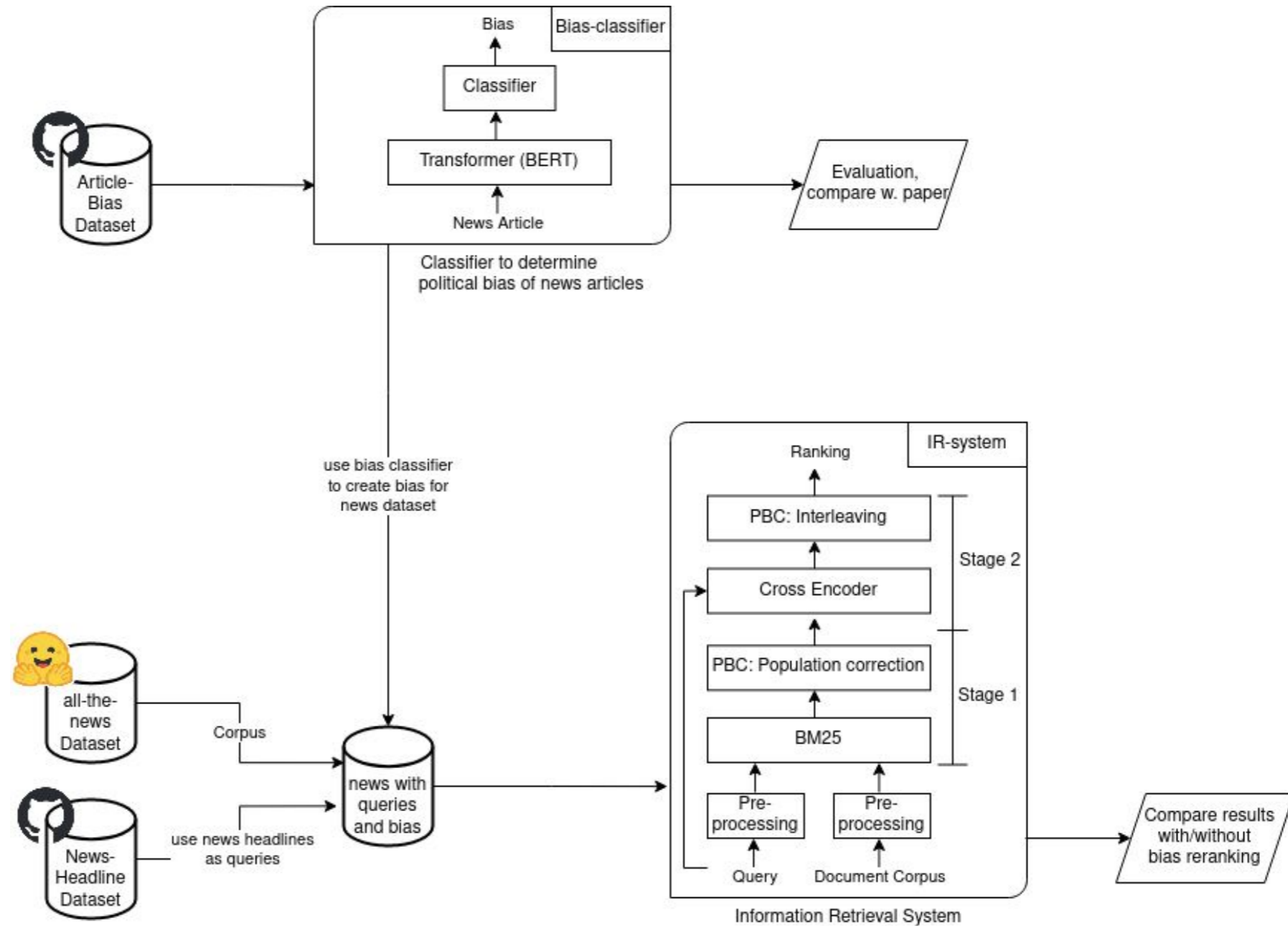## NewSHead dataset


```json
{
    "urls": [
        "https://www.denverpost.com/2019/05/18/
        "https://www.coloradoan.com/story/sport
        "https://www.desmoinesregister.com/stor
        "https://www.thegazette.com/iowa-state-
        "https://www.thegazette.com/subject/spo
    ],
    "label": "Colorado state track and field"
},
{
    "urls": [
        "https://www.cbc.ca/radio/thecurrent/th
        "https://www.cbc.ca/news/world/alabama-
        "https://etcanada.com/news/453165/busy-
    ],
    "label": "Alabama legislation on abortion"
},
```
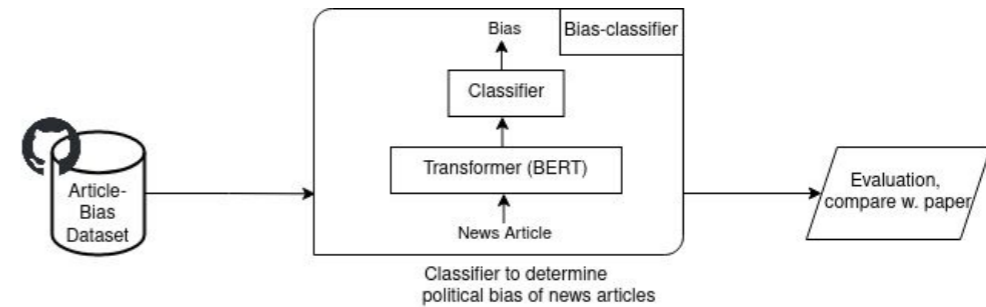
Article-Bias-Prediction

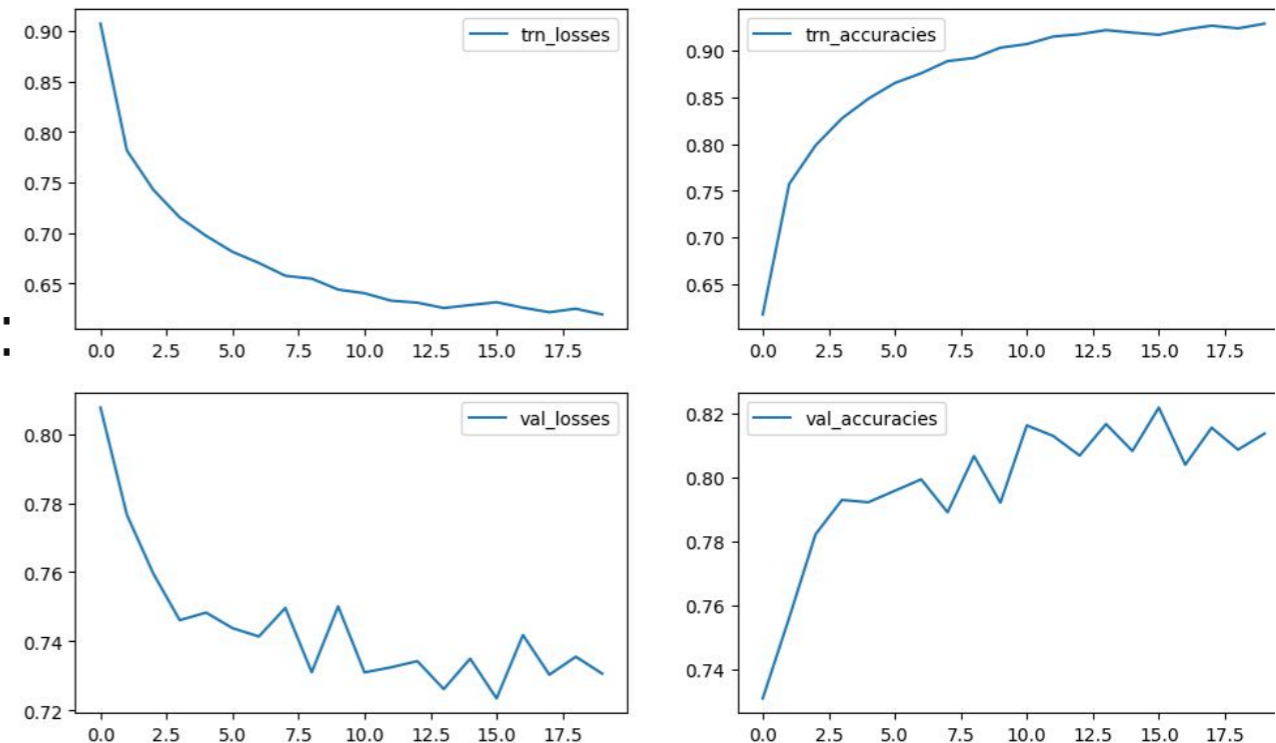| topic	source	bias	url	title	date	authors	content	content_original	source_url	bias_text
environment	NPR Online News	1	http://www.npr.org/blogs/thesalt/2015/04/01/39...	Scary Times For California Farmers As Snowpack...	2015-04-01	Nathan Rott	Scary Times For California Farmers As Snowpack...	Scary Times For California Farmers As Snowpack...	www.npr.org	center
elections	Washington Times	2	http://www.washingtontimes.com/news/2016/jul/4...	Obama joining Clinton on campaign trail to blo...	2016-07-04	Dave Boyer	After restless months on the sidelines , Presi...	After restless months on the sidelines, Presid...	www.washingtontimes.com	right
politics	Guest Writer - Left	0	https://www.cnn.com/2019/03/13/politics/donald...	OPINION: Manafort sentencing highlights Russia...	2019-03-13	Analysis Stephen Collinson	Washington (CNN) Paul Manafort 's final date wi...	Washington (CNN) Paul Manafort's final date wi...	www.cnn.com	left
great_britain	Vox	0	https://www.vox.com/2019/4/11/18305082/brexit-...	The new Brexit deadline is October 31	2019-04-11	Jen Kirby, Hannah Brown, Lauren Katz, Theodore...	European Union leaders agreed in the early mor...	Brexit is delayed, again.\n\nEuropean Union le...	www.vox.com	left
environment	The Verge	0	https://www.theverge.com/2020/5/14/21258650/te...	Tesla's readying a 'million mile' battery that...	2020-05-14	"Sean OKane"	Tesla is getting ready to introduce a long rum...	Tesla is getting ready to introduce a long rum...	www.theverge.com	left

https://huggingface.co/datasets/rjac/all-the-news-2-1-Component-one
https://github.com/google-research-datasets/NewSHead; https://github.com/ramybaly/Article-Bias-Prediction

Overview

Political Bias Classification

Classifier to determine
political bias of news articles

- inspired by *"We Can Detect Your Bias: Predicting the Political Ideology of News Articles"*
- classifier on top of BERT
- classes: left, center, right
- almost reproduced test accuracy: 77% / 79.83%

Training process over 20 epochs

Political Ideology	Count	Percentage
Left	12,003	34.6%
Center	9,743	28.1%
Right	12,991	37.3%

Article-Bias-Prediction dataset

Ramy Baly et al., We Can Detect Your Bias: Predicting the Political Ideology of News Articles

Analysis of IR-Eval Dataset

news with queries and bias
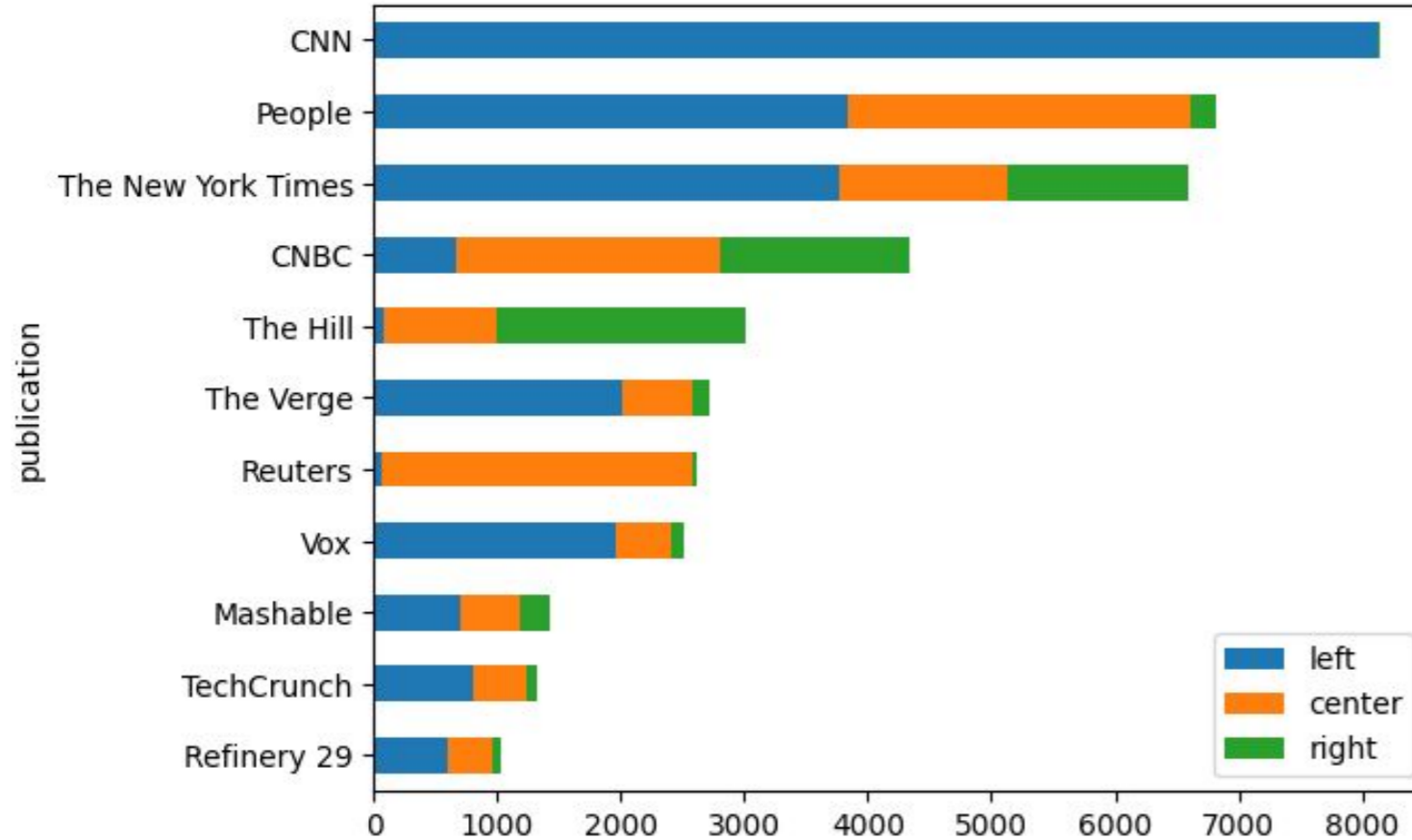
Publisher bias (created using bias-classifier) for outlets with 1000 or more publications
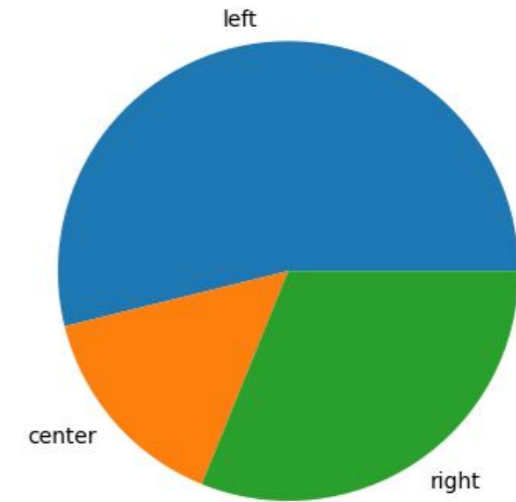
Overall bias (created using bias-classifier)

#relevant docs	how often found
1	29716
2	5844
3	1001
4	114
5	8

Query relevance (created when combining with the NewsSHead dataset)

IR Pipeline

- Simple two-stage IR pipeline
  - S1: Okapi BM25
  - S2: pre-trained cross-encoder
- Political bias correction (PBC) methods in each stage
  - S1: ensure configured minimum docs for each class
  - S2: interleave top `interleave_k` docs by bias class (alternating if possible)

Information Retrieval System
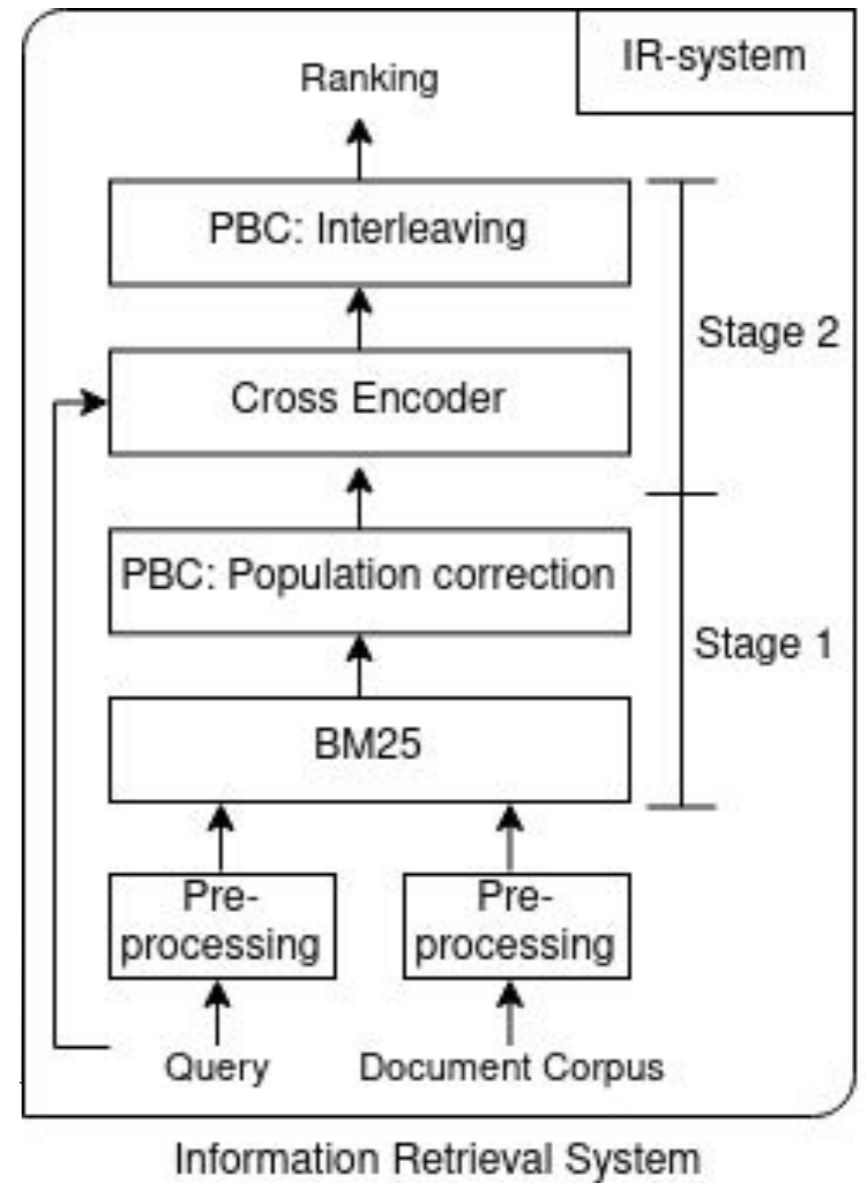
Political Balance Metric for Search Results

- position-sensitive
- class-based (left/center/right)
- normalized
- intuitive
- probability-based

Political Balance Metric for Search Results

(proposed new metric)

- position-sensitive
- class-based (left/center/right)
- normalized
- intuitive
- probability-based
  (but we used one-hot encoded probs in the eval pipeline)

$k \ldots$ Number of retrieved documents

$c \ldots$ Class count (usually 3 for 'left', 'center', 'right')

$P_{ij} \ldots$ Probability of document i being in class j

DCG@k (rel) $\ldots$ DCG@k for relevance scores rel $\in \mathbb{R}^k$

$$d_j = \frac{\text{DCG@k}\left((P^T)_j\right)}{\sum_{l=1} \text{DCG@k}\left((P^T)_l\right)} \ldots \text{Class representation distribution } \vec{d}$$

$$\vec{d} \in \mathbb{R}^c, \sum_{j=1}^{c} d_j = 1$$

$$b = 1 - \frac{\widehat{\text{Var}}\left(\vec{d}\right)}{\underbrace{\widehat{\text{Var}}\left(\left[1, \underbrace{0, \ldots, 0}_{c-1 \text{ times}}\right]\right)}_{\text{worst-case variance}}} \ldots \text{Class Balance in Rankings, } b \in [0,1]$$

b≈0.999 for [l, r, c, c, r]

b=0　　　for [l, l, l, l, l]

b≈0.34　for [l, l, l, l, r]

b≈0.66　for [l, l, l, c, r]

b≈0.96　for [c, r, l, l, l]

Evaluation Results (mean values, rounded to 3 decimals, N=1000 queries)

k = 50, PBC `stage_1_min_fraction_per_class` = 20%, PBC `stage_2_interleave_k` = 10

```
mean(actual relevant doc count) = 3.1
mean(optimal political balance for optimal IR pipeline) = 0.620
```

	no PBC	full	Δfull	S1	ΔS1	S2	ΔS2	ΔS2 to full
precision@5	0.420	0.384	-0.036	0.420	-0.000	0.384	-0.036	-0.000
recall@5	0.674	0.617	-0.057	0.674	-0.000	0.617	-0.057	-0.000
f1@5	0.516	0.472	-0.044	0.516	-0.000	0.473	-0.044	-0.000
nDCG@5	0.646	0.595	-0.051	0.647	0.001	0.594	-0.052	0.001
political-balance@5	0.661	0.859	0.198	0.662	0.002	0.853	0.192	0.005
masked-political-balance@5	0.469	0.481	0.012	0.469	-0.001	0.479	0.010	0.002
precision@20	0.137	0.137	-0.000	0.137	-0.000	0.137	0.000	-0.000
recall@20	0.880	0.878	-0.002	0.878	-0.002	0.880	0.000	-0.002
f1@20	0.237	0.237	-0.001	0.237	-0.001	0.237	0.000	-0.001
nDCG@20	0.740	0.716	-0.024	0.741	0.001	0.715	-0.025	0.001
political-balance@20	0.814	0.854	0.040	0.826	0.012	0.842	0.028	0.012
masked-political-balance@20	0.493	0.507	0.014	0.493	-0.000	0.507	0.014	-0.000

Future work

- **Evaluate** our **proposed metric** based on different scenarios and let some statistics expert verify whether it makes sense
- **Evaluate PBC** methods more intensely
  - using SOTA IR-system
  - using different parameters for PBC methods
  - compare resulting rankings w/ and w/o PBC while ignoring actual relevant docs
  - w/ and w/o PBC on ranking-focused datasets (e.g. some news subset of MS MARCO)
  - use established and peer-reviewed metrics for diversity and popularity in search results
- More advanced PBC methods sensitive to relevance-score for reranking

Conclusion

- Proposed **tunable** methods for political bias correction in rankings
- Proposed intuitive metric for **measuring political balance in search results**
- **Prototype** for **enhancing end-to-end IR pipeline** with PBC methods
  - current results do <u>not</u> show a clear favorable trade-off (!)
    - larger drop in nDCG than rise in masked political balance
    - evaluation on **more suitable dataset** (more relevant documents per query) necessary