pergola_2021_a_disentangled_adversarial_neural_topic_model_for_separating_opinions_from_plots_in_user_reviews

Year

2021

Author(s)

Pergola, Gabriele and Gui, Lin and He, Yulan

Title

A Disentangled Adversarial Neural Topic Model for Separating Opinions from Plots in User Reviews

Venue

NAACL

Topic labeling

Fully automated

Focus

Secondary

Type of contribution

Novel approach?

Underlying technique

Transformer model (Sentence BERT Reimers and Gurevych, 2019

Topic labeling parameters

\

Label generation

The top N terms for each topic fed to Sentence BERT to compute the topic embedding.
The top 10 most similar (annotated) sentences (measured by the cosine similarity) are taken from the dataset.
The most frequent label among the retrieved sentences is adopted as the topic's label.

Motivation

In this context, labels are used as a way to determine how opinion and plot topics are distributed and to compute the topic-disentanglement rate as the proportion of opinion-bearing topics with respect to the overall set of topics, complementary to the proportion of plot/neutral topics.

Topic modeling

Adversarial Neural Topic Model (NTM) (based on Variational Autoencoders Kingma and Welling, 2014)

Topic modeling parameters

- Split of the MOBO dataset into training, development and test set in the proportion of: 80/10/10
- Nr of runs: 5
- Learning rate $\lambda \in [0.001, 0.5]$
- Dropout $\delta \in [0.0, 0.6]$
- Topic vector size $\gamma t \in [25, 50, 100, 200]$
- Encoder and decoder configured following (Srivastava et al., 2017).
- Hidden representation of documents: 100
- Sentiment classifier's hidden size: 50.
- Matrices randomly initialised with the Xavier and sparse methods (Glorot and Bengio, 2010; Martens, 2010).
- Optimiser: Adam (Kingma and Ba, 2015)
- Batch size: 64
- Additional regulariser: Batch normalisation (Cooijmans et al., 2017).

Nr. of topics

Model tested on: 25/50/100/200 topics

Label

Proxy label (Positive, Negative, Plot, None)

Label selection

By majority voting among the labels of the top 10 most similar (annotated) sentences (measured by the cosine similarity) taken from the dataset.

Label quality evaluation

\

Assessors

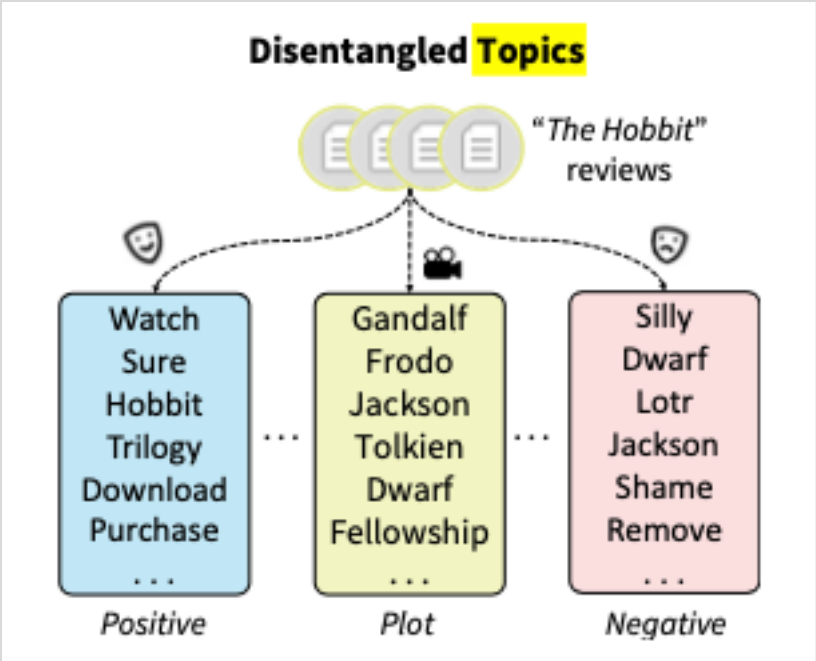
\

Domain

Paper: Topic disentanglement
Dataset: Entertainment (movie and book reviews)

Problem statement

Proposing an NTM capable of disentangling the inferred topic representations (i.e. generating disentangled topics) to improve their interpretability and discriminative power. Achieving this by distinguishing opinion- bearing topics from plot/neutral ones.



Corpus

MOBO dataset. The documents come from different publicly available datasets:

- Standford's IMDB dataset
- GoodReads reviews dataset
- Amazon reviews dataset (only those having a corresponding plot in the MPST dataset)

15 annotators labeled more than 18,000 reviews' sentences (~ 6000 per corpus) marking

- the sentence polarity (Positive, Negative)
- whether a sentence describes its corresponding movie/book Plot, or none of the above (None)

Statistics	IMDB	GoodReads	Amazon
# plots	1,131	150	100
# reviews	25,836	83,852	32,375
% Pos. reviews	0.46	0.33	0.32
% Neg. reviews	0.54	0.50	0.46
% Neu. reviews	0	0.17	0.22
Training set	20,317	65,816	25,883
Dev. set	2,965	9,007	3,275
Test set	2,554	9,029	3,217
# annotated sent.	6,000	6,000	6,000

Table 1: The MOBO dataset statistics.

Document

A movie or book review, paired with the related plot.
Annotations at the sentence-level to indicate polarity and whether a sentence describes the movie/book plot.

Pre-processing

- Tokenization and sentence splitting with SpaCy.
- Stopwords removal
- Removal of tokens shorter than three characters and those with just digits or punctuation.

```
@inproceedings{pergola_2021_a_disentangled_adversarial_neural_topic_model_for_separating_opinions_from_plots_in_user_reviews,
    title = "A Disentangled Adversarial Neural Topic Model for Separating Opinions from Plots in User Reviews",
    author = "Pergola, Gabriele  and
      Gui, Lin  and
      He, Yulan",
    booktitle = "Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies",
    month = jun,
    year = "2021",
    address = "Online",
    publisher = "Association for Computational Linguistics",
    url = "https://aclanthology.org/2021.naacl-main.228",
    doi = "10.18653/v1/2021.naacl-main.228",
    pages = "2870--2883",
    abstract = "The flexibility of the inference process in Variational Autoencoders (VAEs) has recently led to revising traditional probabilistic topic models giving rise to Neural Topic Models (NTM). Although these approaches have achieved significant results, surprisingly very little work has been done on how to disentangle the latent topics. Existing topic models when applied to reviews may extract topics associated with writers{'} subjective
```

```
opinions mixed with those related to factual descriptions such as plot
summaries in movie and book reviews. It is thus desirable to automatically
separate opinion topics from plot/neutral ones enabling a better
interpretability. In this paper, we propose a neural topic model combined with
adversarial training to disentangle opinion topics from plot and neutral ones.
We conduct an extensive experimental assessment introducing a new collection of
movie and book reviews paired with their plots, namely MOBO dataset, showing an
improved coherence and variety of topics, a consistent disentanglement rate,
and sentiment classification performance superior to other supervised topic
models.",
}


@article{kingma_2013_auto_encoding_variational_bayes,
  abstract = {How can we perform efficient inference and learning in directed
probabilistic models, in the presence of continuous latent variables with
intractable posterior distributions, and large datasets? We introduce a
stochastic variational inference and learning algorithm that scales to large
datasets and, under some mild differentiability conditions, even works in the
intractable case. Our contributions are two-fold. First, we show that a
reparameterization of the variational lower bound yields a lower bound
estimator that can be straightforwardly optimized using standard stochastic
gradient methods. Second, we show that for i.i.d. datasets with continuous
latent variables per datapoint, posterior inference can be made especially
efficient by fitting an approximate inference model (also called a recognition
model) to the intractable posterior using the proposed lower bound estimator.
Theoretical advantages are reflected in experimental results.},
  author = {Diederik P Kingma and Max Welling},
  date-added = {2023-02-28 16:15:00 +0100},
  date-modified = {2023-02-28 16:15:00 +0100},
  eprint = {1312.6114},
  month = {12},
  title = {Auto-Encoding Variational Bayes},
  url = {https://arxiv.org/pdf/1312.6114.pdf},
  year = {2013},
  bdsk-url-1 = {https://arxiv.org/pdf/1312.6114.pdf},
  bdsk-url-2 = {https://arxiv.org/abs/1312.6114}}
```

#Thesis/Papers/Initial