Sentiment Analysis & Text Style Transfer on Semi-Parallel Movie Critic Review Corpora Using NLP Translation Techniques

Datasci 266 - Natural Language Processing

Charlie Glass

Berkeley
UNIVERSITY OF CALIFORNIA

Goal

Dataset: Rotten Tomatoes Critic Reviews: https://www.kaggle.com/datasets/stefanoleone992/rotten-tomatoes-movies-and-critic-reviews-dataset?resource=download

1.1M rows, with a review of a movie, by a particular critic, designated ‘Fresh’ or ‘Rotten’, with a 1-2 sentence headline review

1) Build a sentiment analysis classification model that classifies a reviews as ‘Fresh’ or ‘Rotten’
2) Use an Encoder-Decoder Architecture to translate ‘Fresh’ reviews into ‘Rotten’ reviews
    a) Use the sentiment analysis model to evaluate the sentiment of the decoded sequences
3) Use an Encoder-Decoder Architecture to translate the movie critic Roger Ebert’s reviews into a generic review

Challenges

- Only review headline, not full review
- Translation pairs are **non/semi-parallel corpora**
  - **Parallel corpora:** "*I liked the movie*" → "*I hated the movie*"
  - **Non/semi-parallel corpora**: "*After 'Thor,' this makes Marvel Comics two-for-two so far this summer movie season*" → "*Imagine how disappointed I am to find this movie to be a scattered affair that loses power the more it tries to take on big issues*"
- Movie critics have extensive vocabularies
- Reviews can contain difficult cultural phrases/cues

Results

Berkeley
UNIVERSITY OF CALIFORNIA

Sentiment Analysis Model

Model Description	Test Accuracy
Baseline	64%
Bag-of-words Logistic Regression, 500 word vocab	70%
Bag-of-words Neural Network, size-15 hidden layer, 500 word vocab	70%
Bag-of-words Neural Network, size-15 hidden layer, 1000 word vocab	73%
DAN model BERT embeddings	74%
CNN model BERT embeddings	81%
BERT pre-trained model with CNN	88%

Sentiment Transfer Model

Model	About a movie (mean, std)	Contains relevant info (mean, std)	Fluidity Score 1-10 (mean, std)
Baseline	(1.00 , 0.00)	(0.04, 0.20)	(7.06, 0.43)
Negative Sentiment Transfer (20 Epochs)	(1.00 , 0.00)	(0.05, 0.23)	(3.50, 2.02)
Negative Sentiment Transfer (40 Epochs)	(1.00 , 0.00)	(0.14, 0.36)	(4.18, 2.16)
Negative Sentiment Transfer (<85 characters, repeat pairs, 10 epochs)	(1.00 , 0.00)	(0.02, 0.15)	(9.38, 1.94)
Negative Sentiment Transfer (matched on close lengths, 40 epochs)	(1.00 , 0.00)	(0.06, 0.25)	(2.97, 2.11)
Negative Sentiment Transfer (matched on close lengths, 25 epochs)	(1.00 , 0.00)	(0.00, 0.00)	(2.97, 1.72)
Negative Sentiment Transfer (100,000 Vocab Size, 25 epochs)	(1.00 , 0.00)	(0.00, 0.00)	(3.04, 1.00)

***human judges scoring**

***classification model found sentiment of outputs to be negative 83%-95% of the time, across models**

Sentiment Transfer Model

Example Sentences:

"*The film's second half, however, disappoints in many respects.*"

"The film's biggest problem is that the film is, in the end, and the film's lack of a story that is so much more than a a film that is a bit too much, and the film is so predictable."

Ebert Transfer Model

Model	Sentiment matches (mean, std)	About a movie (mean, std)	Contains relevant info (mean, std)	Fluidity Score 1-10 (mean, std)
Baseline	(0.96, 0.07)	(1.0, 0.0)	(0.05, 0.09)	(7.43, 0.78)
Ebert Seq2Seq Transfer Model	(0.76, 0.43)	(1.0, 0.0)	(0.14, 0.35)	(3.33, 2.02)

***human judges scoring**

Ebert Transfer Model

Example Sentences:

"*The wrestler is a film about the human spirit, and the best of the year.*"

"A film that is not only a great film, but also a little too long and too long and too often a little too long, and too much of the film is a little too much."

Conclusion/Next Steps

- For a model to effectively transfer text style with non/semi-parallel corpora, there needs to be more data (more samples, longer reviews), more translation examples, and a more complex architecture than Keras Transformer Encoder/Decoder with own subword model and vocabulary
- Try BERT embeddings/pre-trained model for better word meaning/variance and Cross-Aligned Auto-Encoder Architecture (Shen et al., https://arxiv.org/pdf/1705.09655.pdf) for text style transfer
    - Attempted these during modeling, but RAM and GPU crashed

Thank You!

Questions/ Comments

charliedglass@berkeley.edu