NNFL RESEARCH PAPER ASSIGNMENT

Mohit Dhawan 2017A3PS0223P
Aayush Dandi 2017A3PS0250P
Jatin 2018A3PS0594P

Paper Title

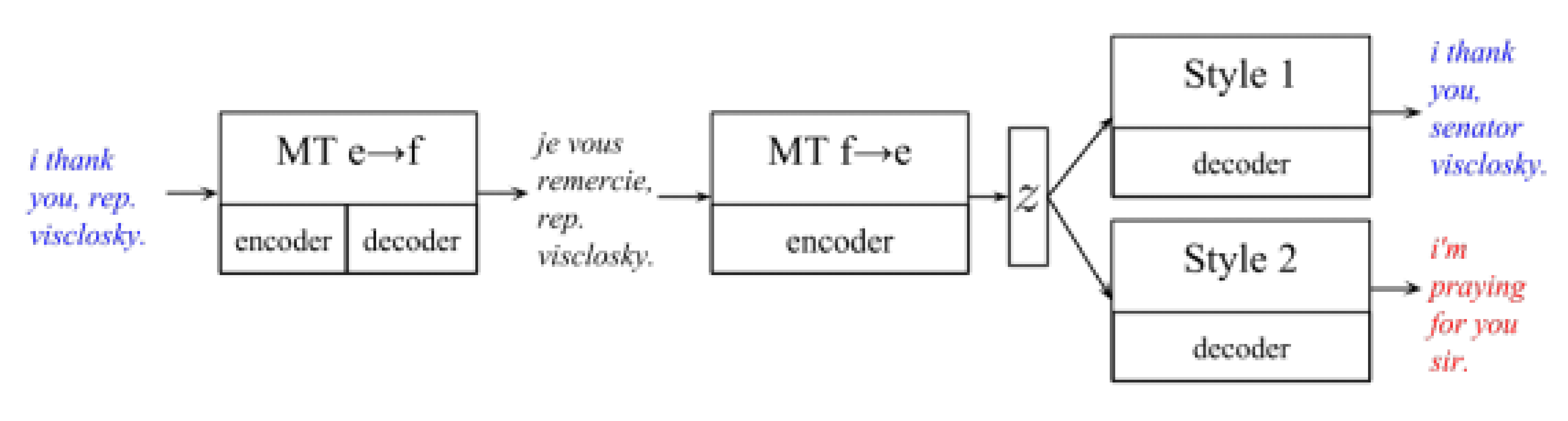
Style Transfer Through Back-Translation[1]

Shrimai Prabhumoye, Yulia Tsvetkov, Ruslan Salakhutdinov, Alan W Black
Carnegie Mellon University, Pittsburgh, PA, USA

Aim

- To accomplish the task of style transfer using machine translation and generation models.

Methodology



- Input datasets are taken as X_1 and X_2 , with styles s_1 and s_2 respectively.
- Input X_i is given to Machine translation model e -> f, and then taken out from machine translation model f -> e, as *z.*

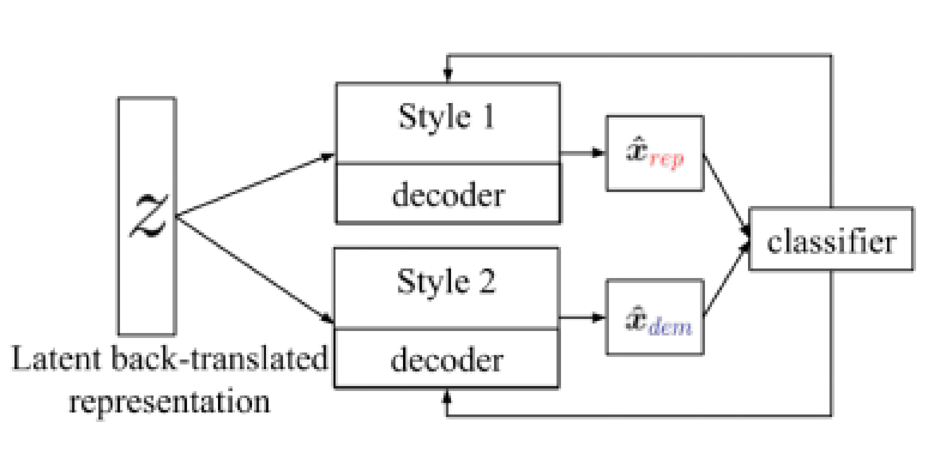
- This is done to accomplish 2 purposes
  - Represent the meaning of the input sentence grounded in back-translation in z
  - Weaken the style attributes of author’s traits

- The reasoning behind this that, as per prior work[2] the process of translating a sentence from a source language to a target language retains the meaning of the sentence but does not preserve the stylistic features related to the author’s traits.

- The machine translation and the encoder of the back-translation model remain fixed. They are not dependent on the data we use across different tasks. This facilitates reusability and spares the need of learning separate models to generate z for a new style data

Generators

- Using the encoder embedding z, we train multiple decoders for each style. The sentence generated by a decoder is passed through the classifier. The loss of the classifier for the generated sentence is used as feedback to guide the decoder for the generation process.



Block diagram of generator

$$\mathcal{L}_{recon}(\theta_G; x) = \mathbb{E}_{q_E(z|x)}[\log p_{gen}(x|z)]$$

Reconstruction loss function

$$\begin{aligned} \hat{x} \sim z &= p(\hat{x}|z) \\ &= \prod_t p(\hat{x}_t|\hat{x}^{<t}, z) \end{aligned}$$

Sample generation equation

$$z = E(x) = q_E(z|x)$$

Latent representation z

$$\min_{\theta_{gen}} \mathcal{L}_{gen} = \mathcal{L}_{recon} + \lambda_c \mathcal{L}_{class}$$

Generative loss L_{gen}

Style Classifier

- A convolutional neural network (CNN) classifier is trained to accurately predict the given style. It is also used to evaluate the error in the generated samples for the desired style. The classifier is trained in a supervised manner, for this labelled examples are given. Θ_c is the tensor of parameters of the classifier.
- To improve the accuracy of the classifier, classifier's inputs with are augmented style-specific lexicons. If a word is present in style-specific lexicon it's prefixed with 1, else it is prefixed with 0. Style lexicons are extracted using the log-odds ratio informative Dirichlet prior (Monroe et al., 2008[3]), a method that identifies words that are statistically overrepresented in each of the categories.

$$\mathcal{L}_{class}(\boldsymbol{\theta}_C) = \mathbb{E}_{\boldsymbol{X}}[\log q_C(\boldsymbol{s}|\boldsymbol{x})].$$

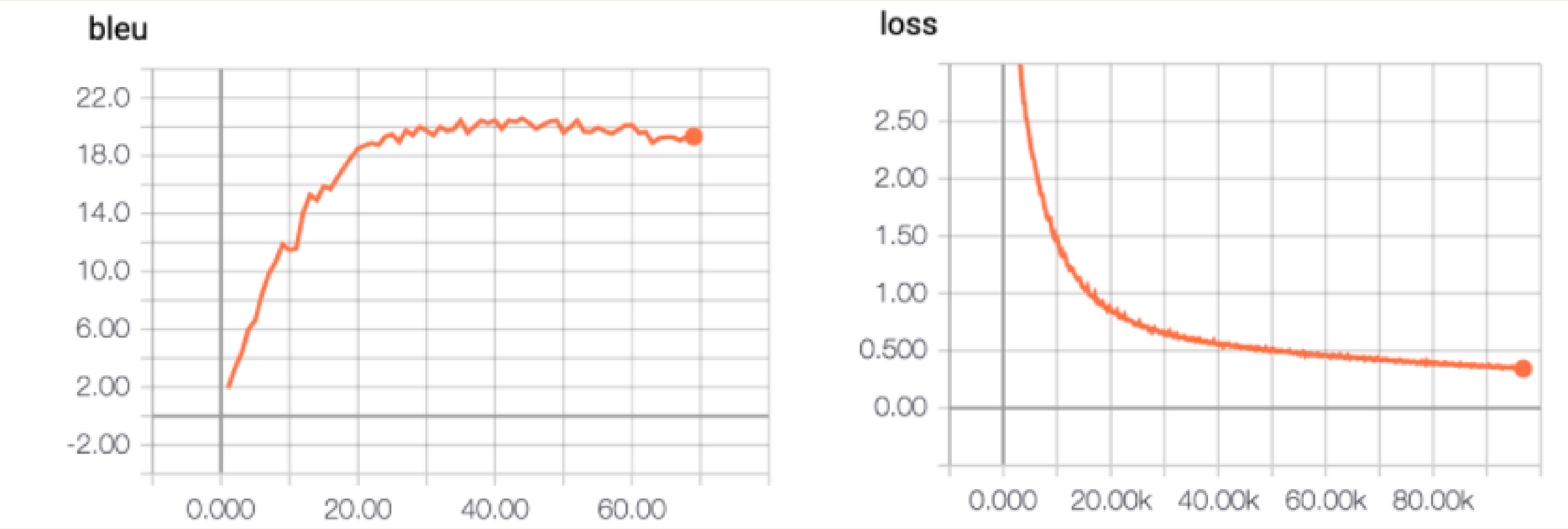
Loss function for classifier (binary cross entropy)

Dataset: Yelp

- The dataset taken for this assignment is the yelp dataset of 1 million plus reviews
- The data is already annotated and hence, classified into positive and negative subsets
- The data has been split into 80:20 for training and testing respectively, the need for a separate validation set was not felt as k-fold cross validation was used during training.

Results

The following were the graphs for bleu score and loss when the model was run on positive to negative style conversion of text.



- Output for negative to positive translation:

Negative (Input)

```
1 ever since joes has it 's just gotten and
2 there is room in that part of the venue .
3 so basically tasted down .
4 'd be back and for a few .
5 how this pharmacy is .
6 just took it .
7 it is n't , but it is
8 i could my birthday gift !
9 new owner , i heard - know the details .
10 but it probably
11 we sit down and we got service .
12 did include miso soup and a small salad .
```

Positive (Output)

```
1 best customer service ive ever been here its just how bad and it 's always delicious .
2 there is also a good room in that part of the venue .
3 so basically tasted watered down .
4 be a great place to come and try without a few .
5 how great customer service is this pharmacy is .
6 just took it .
7 it is n't bad , but it 's always good and the service is great .
8 i could definitely make my birthday gift !
9 new owner , i know its nice and friendly and know the details of the animals .
10 but it probably definitely be my favorite !
11 we sit down and we got great service .
12 did a great job include broccoli and a decent beer with generous portions .
```

Challenges faced

- Due to limited computational resources, we couldn't experiment on as many datasets as we would have liked
- Due to limited time and other course work we couldn't expand the project to as many areas as we would've liked.

Experience of assignment and learning outcomes

- We learnt about bi-directional sequential models, use encodings in nlp.
- It was great to collaborate virtual with people of similar enthusiasm, especially in these trying times.
- We came to know about the wide range of research that is going on in the field of text generation.
- We got to know better about github and its usage.

References

- [1] Prabhumoye, S., Tsvetkov, Y., Salakhutdinov, R., & Black, A. W. (2018). Style transfer through back-translation. *arXiv preprint arXiv:1804.09000*.
- [2] Ella Rabinovich, Shachar Mirkin, Raj Nath Patel, Lucia Specia, and Shuly Wintner. 2016. Personalized machine translation: Preserving original author traits. In Proc. EACL.
- [3] Burt L. Monroe, Michael P. Colaresi, and Kevin M. Quinn. 2008. Fightin words: Lexical feature selection and evaluation for identifying the content of political conflict. Political Analysis

Thank you