

Stylized Text Generation: Approaches and Applications

Lili Mou

University of Alberta; Amii
doublepower.mou@gmail.com

Olga Vechtomova

University of Waterloo
ovechtom@uwaterloo.ca

Type of Tutorial: Cutting-edge.

1 Tutorial Introduction

Text generation has played an important role in various applications of natural language processing (NLP), such as paraphrasing, summarization, and dialogue systems. With the development of modern deep learning techniques, text generation is usually accomplished by a neural decoder (e.g., a recurrent neural network or a Transformer), which generates a word at a time conditioned on previous generated words. The decoder can be further conditioned on some source information, such as a source language sentence in machine translation, or a previous utterance in dialogue systems.

In recent studies, researchers are paying increasing attention to modeling and manipulating the style of the generation text, which we call *stylized text generation* in this tutorial. The goal is to not only model the content of text (in traditional text generation), but also control some “style” of the text, for example, the persona of a speaker in a dialogue (Li et al., 2016), or the sentiment of product reviews (Hu et al., 2017).

Stylized text generation is related to various machine learning techniques, for example, embedding learning techniques to represent style (Fu et al., 2018), adversarial learning and reinforcement learning with cycle consistency to match “content” but to distinguish different styles (Hu et al., 2017; Xu et al., 2018; John et al., 2019); very recent work is even able to disentangle latent features in an unsupervised way (Xu et al., 2019).

In this tutorial, we will provide a comprehensive literature review on stylized text generation. We start from the definition of style and different settings of stylized text generation, illustrated with various applications.

In the second part, we will describe style-

conditioned text generation. In this category, style serves as a certain type of source information, which the decoder is conditioned on. We describe three types of approaches: (1) embedding-based techniques that capture the style information by real-valued vectors, which can be used to condition a language model (Tikhonov and Yamshchikov, 2018) or concatenated with the input to a decoder (Li et al., 2016; Vechtomova et al., 2018) (2) approaches that encode both style and content in the latent space (Shi et al., 2019a; Yang et al., 2017; Li et al., 2020). We will discuss techniques that structure latent space to encode both style and content, and include Gaussian Mixture Model Variational Autoencoders (GMM-VAE) (Shi et al., 2019a; Wang et al., 2019a; Shi et al., 2019b), Conditional Variational Autoencoders (CVAE) (Yang et al., 2017), and Adversarially Regularized Autoencoders (ARAE) (Li et al., 2020). (3) approaches with multiple style-specific decoders (Syed et al., 2019; Chen et al., 2019). We highlight several applications including persona-based dialogue generation (Li et al., 2016) and creative writing (Yang et al., 2017; Tikhonov and Yamshchikov, 2018; Vechtomova et al., 2018).

Next, we will introduce evaluation methods for style-conditioned text generation. We will present the current practice in the literature, involving both human evaluation and automatic metrics. A few important evaluation aspects include the success of being in the target style, the preservation of content information, as well as language fluency in general.

In the third part, we will focus on style-transfer text generation. Given an input sentence of a certain style, the goal of style transfer is to synthesize a new sentence that has the same content but with different styles. Particularly, style-transfer text generation can be categorized into three settings: (1) Parallel-supervised style transfer, where a par-

allel corpus is available (Xu et al., 2012; Rao and Tetreault, 2018). This is similar to machine translation, but semi-supervised learning is adopted to address small-data training (Wang et al., 2019b). (2) Non-parallel style transfer, where each sentence is annotated by a style label (e.g., positive or negative sentiment). This setting is the most explored setting in previous style transfer literature. We will discuss classification losses to distinguish different styles (John et al., 2019), and adversarial losses/cycle consistency to match content information (Shen et al., 2017). We will also present an editing-based approach that edits style-specific words and phrases into the desired style (Li et al., 2018). (3) Unsupervised style transfer, where the entire corpus is unlabeled (no parallel pairs or style labels). In recent studies, researchers have applied auxiliary losses (such as orthogonality penalty) to detect the most prevalent variation of text in a corpus, and are sometimes able to accomplish style transfer in a purely unsupervised fashion. Since unsupervised style transfer is new to NLP and less explored, we will also introduce several studies in the computer vision domain, bringing future opportunities to text generation in this setting (Gatys et al., 2016; Chen et al., 2016).

Next, we will discuss style adversarial text generation (Zhang et al., 2019). The setting of adversarial attacks is similar to style transfer in that it aims to change the style classifier’s prediction. However, the synthesized sentence in this setting should in fact keep the actual style as humans perceive, but “fool” the style classifier. Thus, it is known as the *adversarial attack*. We will discuss style adversarial generation in the character level, the word level, as well as the sentence level. Techniques include discrete word manipulation and continuous latent space manipulation.

Finally, we will conclude our tutorial by presenting the challenges of stylized text generation and discussing future directions, such as small-data training, non-categorical style modeling, and a generalized scope of style transfer (e.g., controlling the syntax as a style (Bao et al., 2019)).

By the end of the tutorial, the audience will have a systematic view of different settings of stylized text generation, understand common techniques to model and manipulate the style of text, and be able to apply existing approaches to new scenarios that require stylized text generation. Our tuto-

rial also investigates stylized generative models in non-NLP domains, and thus would inspire future NLP studies in this direction.

2 Tutorial Outline

PART I: Introduction (20 min)

- Definition of style
- Settings and Problem formulations
- Examples of style (e.g., sentiment, artistic style, grammatical style)

PART II: Style-Conditioned Text Generation (50 min)

- Techniques
 - Encoding style in embeddings: sequence-to-sequence models with style embeddings, style conditioned language models, Variational Autoencoder (VAE) with style embeddings;
 - Encoding style and content in latent space: Conditional Variational Autoencoder (CVAE) Gaussian Mixture Variational Autoencoder (GMM-VAE), Adversarially Regularized Autoencoder (ARAE).
 - Models with multiple style-specific decoders
- Applications
 - Creative text generation (e.g., poetry composition)
 - Persona and emotion conditioned dialogue models
 - Stylized image caption generation
- Evaluation measures
 - Stylistic adherence
 - Content preservation
 - Language fluency
 - Novelty and diversity

PART III: Style-Transfer Text Generation (60 min)

- Parallel supervised style transfer
 - Sequence-to-sequence learning
 - Semi-supervised training with limited parallel data

- Applications: Shakespearean–modern English transfer, formality style transfer
- Non-parallel supervised style transfer
 - Auxiliary classification for style modeling
 - Adversarial learning for matching content
 - Cycle consistency for content matching
 - Edit-based style transfer
 - Applications: Sentiment, genre and grammatical style transfer
- Unsupervised style transfer
 - Approaches: Mutual information penalties and correlation penalties for automatic style detection
 - A brief introduction of unsupervised style transfer in image domain (e.g., color, shape, angle)

PART IV: Style-Adversarial Text Generation
(30 minutes)

- Style adversarial vs. style transfer
- Approaches
 - Character-level attack
 - Word-level attack
 - Sentence-level attack

PART IV: Conclusion, Future Work, and Q&A
(20 min)

- Challenges: non-categorical style, small-data training
- A broader view of “style”: text summarization/simplification as style transfer, syntax-semantic disentanglement

3 Instructors

Lili Mou

doublepower.mou@gmail.com
<https://lili-mou.github.io>

Dr. Lili Mou is an Assistant Professor at the Department of Computing Science, University of Alberta. He is also an Amii Fellow and a Canadian CIFAR AI Chair. Lili received his BS and PhD degrees in 2012 and 2017, respectively, from School of EECS, Peking University. After that, he worked as a postdoctoral fellow at the University of Waterloo and a research scientist at Adeptmind (a startup in Toronto, Canada). His research

interests include deep learning applied to natural language processing as well as programming language processing. Recently, he has been focusing more on text generation, from both continuous latent space and discrete word space. He has more than 30 papers published at top-tier conferences and journals, including AAAI, ACL, CIKM, COLING, EMNLP, ICASSP, ICLR, ICML, IJCAI, INTERSPEECH, NAACL-HLT, and TACL. He presented a tutorial “*Discreteness in Neural Natural Language Processing*” at EMNLP-IJCNLP’19.

Olga Vechtomova

ovechtomova@uwaterloo.ca
<https://ov-research.uwaterloo.ca>

Dr. Olga Vechtomova is an Associate Professor in the Department of Management Sciences, Faculty of Engineering, cross-appointed in the School of Computer Science at the University of Waterloo. Olga leads the Natural Language Processing Lab, affiliated with the Waterloo.AI Institute. Her research has been supported by a number of industry and government grants, including Amazon Research Award and Natural Sciences and Engineering Research Council (NSERC). The research in her Lab is mainly focused on designing deep neural networks for natural language generation tasks. Her current and recent projects include controlled text generation, text style transfer, and designing text generative models for creative applications. She has over 50 publications in NLP and Information Retrieval conferences and journals, including NAACL-HLT, COLING, ACL, ACM SIGIR, and CIKM. She and her colleagues recently received the ACM SIGIR 2019 Test of Time Award.

Acknowledgments

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), under grant Nos. RGPIN-2019-04897 and RGPIN-2020-04465. Lili Mou is also supported by AltaML, the Amii Fellow Program, and the Canadian CIFAR AI Chair Program.

References

Yu Bao, Hao Zhou, Shujian Huang, Lei Li, Lili Mou, Olga Vechtomova, Xin-yu Dai, and Jiajun Chen. 2019. [Generating sentences from disentangled syntactic and semantic spaces](#). In *ACL*, pages 6008–6019.

- Cheng-Kuan Chen, Zhufeng Pan, Ming-Yu Liu, and Min Sun. 2019. Unsupervised stylish image description generation via domain layer norm. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 8151–8158.
- Xi Chen, Yan Duan, Rein Houthoofd, John Schulman, Ilya Sutskever, and Pieter Abbeel. 2016. [InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets](#). In *NIPS*, pages 2172–2180.
- Zhenxin Fu, Xiaoye Tan, Nanyun Peng, Dongyan Zhao, and Rui Yan. 2018. [Style transfer in text: Exploration and evaluation](#). In *AAAI*, pages 663–670.
- Leon A Gatys, Alexander S Ecker, and Matthias Bethge. 2016. [Image style transfer using convolutional neural networks](#). In *CVPR*, pages 2414–2423.
- Zhiting Hu, Zichao Yang, Xiaodan Liang, Ruslan Salakhutdinov, and Eric P Xing. 2017. [Toward controlled generation of text](#). In *ICML*, pages 1587–1596.
- Vineet John, Lili Mou, Hareesh Bahuleyan, and Olga Vechtomova. 2019. [Disentangled representation learning for non-parallel text style transfer](#). In *ACL*, pages 424–434.
- Jiwei Li, Michel Galley, Chris Brockett, Georgios Spathourakis, Jianfeng Gao, and Bill Dolan. 2016. [A persona-based neural conversation model](#). In *ACL*, pages 994–1003.
- Juncen Li, Robin Jia, He He, and Percy Liang. 2018. [Delete, retrieve, generate: a simple approach to sentiment and style transfer](#). In *NAACL-HLT*, pages 1865–1874.
- Yuan Li, Chunyuan Li, Yizhe Zhang, Xiujuan Li, Guoqing Zheng, Lawrence Carin, and Jianfeng Gao. 2020. Complementary auxiliary classifiers for label-conditional text generation. In *AAAI*.
- Sudha Rao and Joel Tetreault. 2018. [Dear sir or madam, may I introduce the GYAFC dataset: Corpus, benchmarks and metrics for formality style transfer](#). In *NAACL-HLT*, pages 129–140.
- Tianxiao Shen, Tao Lei, Regina Barzilay, and Tommi Jaakkola. 2017. [Style transfer from non-parallel text by cross-alignment](#). In *NIPS*, pages 6830–6841.
- Wenxian Shi, Hao Zhou, Ning Miao, Shenjian Zhao, and Lei Li. 2019a. [Fixing Gaussian mixture VAEs for interpretable text generation](#). *arXiv preprint arXiv:1906.06719*.
- Wenxian Shi, Hao Zhou, Ning Miao, Shenjian Zhao, and Lei Li. 2019b. [Fixing Gaussian mixture VAEs for interpretable text generation](#). *arXiv preprint arXiv:1906.06719*.
- Bakhtiyar Syed, Gaurav Verma, Balaji Vasan Srinivasan, Vasudeva Varma, et al. 2019. [Adapting language models for non-parallel author-stylized rewriting](#). *arXiv preprint arXiv:1909.09962*.
- Alexey Tikhonov and Ivan P Yamshchikov. 2018. [Guess who? multilingual approach for the automated generation of author-stylized poetry](#). In *IEEE Spoken Language Technology Workshop*, pages 787–794.
- Olga Vechtomova, Hareesh Bahuleyan, Amirpasha Ghabussi, and Vineet John. 2018. [Generating lyrics with variational autoencoder and multi-modal artist embeddings](#). *arXiv preprint arXiv:1812.08318*.
- Wenlin Wang, Zhe Gan, Hongteng Xu, Ruiyi Zhang, Guoyin Wang, Dinghan Shen, Changyou Chen, and Lawrence Carin. 2019a. [Topic-guided variational autoencoders for text generation](#). *arXiv preprint arXiv:1903.07137*.
- Yunli Wang, Yu Wu, Lili Mou, Zhoujun Li, and Wenhao Chao. 2019b. [Harnessing pre-trained neural networks with rules for formality style transfer](#). In *EMNLP-IJCNLP*.
- Jingjing Xu, Xu Sun, Qi Zeng, Xiaodong Zhang, Xuancheng Ren, Houfeng Wang, and Wenjie Li. 2018. [Unpaired sentiment-to-sentiment translation: A cycled reinforcement learning approach](#). In *ACL*, pages 979–988.
- Peng Xu, Yanshuai Cao, and Jackie Chi Kit Cheung. 2019. [Unsupervised controllable text generation with global variation discovery and disentanglement](#). *arXiv preprint arXiv:1905.11975*.
- Wei Xu, Alan Ritter, Bill Dolan, Ralph Grishman, and Colin Cherry. 2012. [Paraphrasing for style](#). In *COLING*, pages 2899–2914.
- Xiaopeng Yang, Xiaowen Lin, Shunda Suo, and Ming Li. 2017. [Generating thematic chinese poetry using conditional variational autoencoders with hybrid decoders](#). *arXiv preprint arXiv:1711.07632*.
- Huangzhao Zhang, Hao Zhou, Ning Miao, and Lei Li. 2019. [Generating fluent adversarial examples for natural languages](#). In *ACL*, pages 5564–5569, Florence, Italy.