STYLE EXAMPLE-GUIDED TEXT GENERATION USING GENERATIVE ADVERSARIAL TRANSFORMERS

Anonymous authors
Paper under double-blind review

ABSTRACT

We introduce a language generative model framework for generating a styled paragraph based on a context sentence and a style reference example. The framework consists of a style encoder and a texts decoder. The style encoder extracts a style code from the reference example, and the text decoder generates texts based on the style code and the context. We propose a novel objective function to train our framework. We also investigate different network design choices. We conduct extensive experimental validation with comparison to strong baselines to validate the effectiveness of the proposed framework using a newly collected dataset with diverse text styles. Both code and dataset will be released upon publication.

1 INTRODUCTION

Text generation refers to the task of generating a new text based on some user input. The task exists in many forms, but arguably the most common form concerns generating a coherent and consistent text based on an input context such as the first few sentences of the target output. This is often achieved by giving the context to a generative language model. Generative language models play a central role in machine learning and natural language processing (NLP). Not only they serve as the main mean for unsupervised feature representation learning but also find use in various applications, including question answering, dialogue agents, summarization, and content creation systems.

Thanks to the introduction of novel deep learning architectures and the availability of large-scale training corpora, the state-of-the-art text generation has advanced significantly in recent years. We can now train language models capable of generating fluent and coherent texts that people cannot tell them apart from those written by humans. However, despite the great achievement, existing generative models are limited and inflexible in the sense that a trained model is only capable in generating texts of one style. It can not be used to generate texts of different styles. For instance, a news generative model can only be used to generate news, and a lyric generative model can only be used to generate lyrics. In contrast, humans can compose texts in various styles.

To bridge the gap, we propose a style example-guided text generation framework that can generate styled texts based on the style of the example reference text. In our framework, the generator takes two inputs where one is the context input while the other is the style reference example. We use the style reference example to change the generation behavior of our generative model dynamically. For a fixed context, when the provided style reference is a paragraph sampled from a news article, it becomes a news generator. When the provided style reference is a review, it becomes a review generator. In other words, the same generator can generate texts of different styles based on the examples. In Tab.1 we show example outputs of the proposed framework where we generate texts of dramatically different styles for the same input sentence.

The proposed style example-guided text generation framework is based on the generative adversarial networks (GANs), and we utilize the transformer in both the generator and discriminator design. We collect a large dataset containing documents with many different styles for training. Using a novel learning objective function, our network learns to generate styled texts based on the input style example in an unsupervised manner. We conduct extensive experimental validations with comparisons to strong baselines. We also investigate different ways of designing the generator and compare their performance. Through detailed quantitative and user study results, we prove the effectiveness of the proposed framework for the style example-guided text generation task.
Table 1: Given the same context sentences but different style reference examples, our generation model can generate paragraphs of different styles. The output paragraph shares the same style as the reference example.

## 2 Related Work

**Language modeling** has seen many advancements in recent years, which has resulted in significant improvements on various NLP tasks. Early language models focused on using n-grams to represent a text distribution. Bengio et al. (2003) introduced a neural language model in a shift from more traditional n-gram models. Many works later (Mikolov et al. (2013); Pennington et al. (2014)) focused on word embeddings as a way to represent tokens within the text. More recently, Peters et al. (2018) used bi-directional LSTMs to obtain deep contextualized word representation. However, RNNs can only represent a limited context. Vaswani et al. (2017) introduced the transformer networks which use the connections between long-distance word pairs embedded in attention mechanisms and can easily enable the learning of long-term dependency. Many later models (Devlin et al. (2019); Liu et al. (2019); Dai et al. (2019); Yang et al. (2019)) used transformer model and obtained significant improvements on downstream tasks (Wang et al. (2019); Rajpurkar et al. (2016); Zellers et al. (2018)). Lately, Radford et al. (2019) introduced GPT-2, a generative left-to-right language model based on the transformer and showed that these models are able to generate coherent text when pre-trained on a large corpus. Shoeybi et al. (2019) further scaled up the GPT-2 model and demonstrated improved performance. Our work differs from the prior works because we aim for allowing user flexible control over the style of the generated text.

**Texts generation** includes review generation (Radford et al. (2018); Zang & Wan (2017)), sentiment texts generation (Wang & Wan (2018); Hu et al. (2017); Merity et al. (2017)), Wikipedia generation (Liu et al. (2018); Lebret et al. (2016)), fake news generation (Bahktin et al. (2019); Zellers et al. (2019)), abstractive summarization (Li et al. (2018); Zhang et al. (2019); Pasunuru et al. (2017)), and conversation/dialogue system (Vinyls & Le (2015); Budzianowski & Vulic (2019)). Although many of them trained a transformer on large-scale corpora, their results were limited in their specific domain (e.g., reviews, news, etc.) because they either utilized domain-specific priors in their model design or were not designed to generate texts in many different domains or styles.

**Control on texts generation.** In addition, there are literature utilizing insertion-base (Stern et al. (2019); Chan et al. (2019)), GAN-based (Yu et al. (2017); d’Autume et al. (2019)), variational autoencoder-based (Xu et al. (2019)), normalizing flow-based (Tran et al. (2019)) approaches for general texts generation task. However, we instead focus on generating styled paragraphs conditioning on a context and a reference paragraph. A recent work by Keskar et al. (2019) is most related to ours. They propose a conditional transformer using a control code to perform language generation in a sequence-to-sequence manner. We demonstrate our method outperforms theirs by a large margin in the experiment section.

**Text style transfer** concerns transferring an input text of one style to a different style (Kerpedjiiev (1992); Rao & Tetreault (2018); Xu (2017); Xu et al. (2012); Fu et al. (2018); Hu et al. (2017);
Our generator, embedding layer, and an output embedding layer. Using a style reference example, given few context sentences, we propose a language generative model framework that allows us to control the style of the output text. Formulation (Achlioptas et al. (2017)), we first divide inputs. To avoid non-differentiability in text decoding (e.g., beam search), we use a latent GAN extracted by text, we use a pretrained trained GPT-2 model. We use a conditional GAN where signals produced by the generator converges to the distribution of signals observed in the real world. In this case, the distribution of the output conditions, the generator learns to convert a random noise vector to a realistic signal in a way that GAN defines a zero-sum game played by a generator and a discriminator \( D \). Under some nice conditions, the generator learns to convert a random noise vector to a realistic signal in a way that the discriminator cannot tell it apart from real signals. In this case, the distribution of the output signals produced by the generator converges to the distribution of signals observed in the real world.

We use a conditional GAN where \( F \) takes a context sentence and a style reference example as inputs. To avoid non-differentiability in text decoding (e.g., beam search), we use a latent GAN formulation (Achlioptas et al. (2017)). We first divide \( F \) into a feature extractor \( F_f \) and an output embedding layer \( F_o \), that is \( F \equiv F_o \circ F_f \). Now, instead of using the output text from \( F_o \) as the discriminator input, we feed the latent representation computed by \( F_f \) to the discriminator. For real text, we use a pretrained trained GPT-2 model \( H \). Again, we decompose \( H \) into a feature extractor \( H_f \) and an output embedding layer \( H_o \) (\( H \equiv H_o \circ H_f \)). The GAN discriminator then takes features extracted by \( H_f \) as input for real texts. Using this latent GAN formulation, we aim for aligning the feature distribution of our generator to the feature distribution of the pretrained GPT-2 model.

### 4 Style Example-Guided Text Generation

We propose a language generative model framework that allows us to control style of the output text using a style reference example. Given few context sentences \( \mathbf{w} = \{w_t\}_{t=1}^{T} \) and a reference text \( s \), our generator \( F \) generates output text \( y \) that has the same style as the reference example \( s \) given by

\[
y = F(\mathbf{w}, s) \equiv F_o(F_f(\mathbf{w}, s)).
\]

We divide the feature extractor \( F_f \) into a style encoder \( F_s \) and a text decoder \( F_g \) where the style encoder extracts a style representation from the style example, \( z = F_s(s) \), and the text decoder \( F_g \) consumes the style representation and the context sentences to compute a feature for \( F_o \) to generate...
the total number of tokens in $p_i$ where $i \sim D$. (a) The reconstruction stream is trained using the language modeling loss $L_{LM}$ and the distillation loss $L_{DIST}$. (b) The cross-style generation stream is trained using the style loss $L_{STYLE}$ and the GAN loss $L_{GAN}$. Note that we decompose each network into a feature extractor and an embedding layer.

the styled text $y$. In this section, we will first introduce the data streams employed during training and our novel learning objective function. We will then discuss various generator design choices.

### 4.1 Learning Data Streams

Let $D = \{(d_n, l_n)\}$ be a dataset of documents where $d_n$ is a document and $l_n$ is its style label. We assume a finite set of style labels $L = \{1, 2, \ldots, L\}$ where each integer represents a style class such as news, review, lyric, poem, novel, and children book. During training, our framework employs two data streams where the first one is called the reconstruction stream while the other is referred to as the cross-style generation stream. We note that such a two-stream processing pipeline is common in GAN-based image translation frameworks [Liu et al. (2017), Huang et al. (2018), Liu et al. (2019a)] but is less explored for language modeling.

**Reconstruction stream** (RS). For this stream, we first sample two documents with the same style from $D$: $(d_i, l_i)$ and $(d_j, l_j)$ where $l_i = l_j$. We then sample two paragraphs $p_i \sim d_i$ and $p_j \sim d_j$. We extract the first few sentences from $p_i$ as the input context $w = \psi(p_i)$, where $\psi$ is the extraction function, and use $p_j$ for the style reference $s$. Feeding $w$ and $p_j$ to the generator $F$, we expect $F$ should be able to reconstruct $p_i$: $F(\psi(p_i), p_j) \approx p_i$.

**Cross-style generation stream** (CS). We first sample two documents $(d_i, l_i) \sim D$ and $(d_k, l_k) \sim D$ where $l_i \neq l_k$. We then sample paragraphs $p_i \sim d_i$ and $p_k \sim d_k$. We again extract the first few sentences from $p_i$ as the input context $w = \psi(p_i)$ and use $p_k$ for the style reference $s$. As feeding $w$ and $p_k$ to the generator $F$, we expect $F$ should output $p_{i \rightarrow k} = F(\psi(p_i), p_k)$ where $p_{i \rightarrow k}$ should has the same style as $d_k$. Let $C^*$ be an oracle style comparator function that outputs 1 if the two input texts have the same style and 0 otherwise. We aim for $C^*(p_{i \rightarrow k}, p_k) = 1$.

### 4.2 Learning Objective

We propose an objective function consisting of four carefully designed loss terms for training the proposed framework using the above two data streams. The objective function is given by

$$ L = L_{LM} + \lambda_{DIST} L_{DIST} + \lambda_{STYLE} L_{STYLE} + \lambda_{GAN} L_{GAN}, $$

where $L_{LM}$ is the language modeling loss, $L_{DIST}$ the distillation loss, $L_{STYLE}$ is a style comparison loss, and $L_{GAN}$ is the latent GAN loss. The scalars $\lambda_{DIST}$, $\lambda_{STYLE}$, and $\lambda_{GAN}$ are the hyper-parameters controlling relative importance of the terms. The values for these hyperparameters and the method for determining their values are discussed in Appendix A. We visualize training with the proposed objective function using the two data streams in Fig. 1.

**Language modeling loss** $L_{LM}$ formulates the probability distribution of a paragraph $p$ as the product of the conditional probability of each token $w_t$ given the previous tokens $\{w_{t-1}\}$ as shown in [1]. We use $L_{LM}$ to supervise the training of the data reconstruction stream. It is given by

$$ L_{LM} = E_{(p_i, p_j) \sim RS} \left[ \frac{1}{T} \sum_{t=1}^{T} \log \left( \frac{e^{F(w_t|w_{t-1}, p_j)}}{\sum_{y} e^{F(w_t|w_{t-1}, p_j)}} \right) \right], $$

where $(p_i, p_j) \sim RS$ denotes that $p_i$ and $p_j$ are from the reconstruction stream. The variable $T$ is the total number of tokens in $p_i$ and $V$ is the size of the vocabulary.

---

1For the purpose of data augmentation, in our implementation, a paragraph we sample may not be the full paragraph in the nominal sense. It could start from the middle of a nominal paragraph.
**Distillation loss.** We use $L_{DIST}$ to regularize the learning as processing the data reconstruction stream. We pretrain a GPT-2 model using our dataset $D$ and use it as our distillation target. We denote the pretrained GPT-2 model as $H$. (Note that $H$ does not have the desired style control capability.) By jointly optimizing $L_{LM}$ and $L_{DIST}$, we train $F$ to generate fluent texts (by minimizing $L_{LM}$) as well as behave similarly to $H$ (by minimizing $L_{DIST}$). The distillation loss is calculated by minimizing the mutual information between output distributions of $F$ and $H$, which is given by

$$L_{DIST} = E_{(p_t, p_j) \sim CS} \left[ -\frac{1}{T} \sum_{t=1}^{T} \sum_{v=|\nu|}^{|\nu|} e^{H(w_{v,t-1}|w_{1:t-1})} \log \left( \frac{e^{F(w_t|w_{1:t-1},p_t)}}{\sum_{v} e^{F(w_v|w_{1:t-1},p_j)}} \right) \right]. \tag{5}$$

We note that the distillation loss has been used in various tasks including model compression, transfer learning, life-long learning, etc (Hinton et al. (2015); Kim & Rush (2016); Liu et al. (2019c); Mirzadeh et al. (2019); Liu et al. (2019b); Hou et al. (2018)). In this paper, we extend its use to the style example-guided language generative model training task.

**Style loss** $L_{STYLE}$ helps ensure the output from the cross-style generation stream has the same style as the input reference. A pretrained style comparator $C$ is used for computing the loss. The comparator takes two paragraphs as input and is trained to output 1 when the two paragraphs have the same style and 0 otherwise. We use $D$ for pretraining $C$ since it contains style labels for each document. We pretrain $C$ using the binary cross-entropy loss. The comparator $C$ is highly accurate. It achieves a classification accuracy of 87.8% to 98.8% in our held-out validation sets. After pretraining, we fix $C$ and use it to train $F$. The style loss $L_{STYLE}$ is then given by

$$L_C = E_{(p_t, p_k) \sim CS} \left[ - \log \left( C(H_f(p_k), F_f(p_i, p_k)) \right) \right], \tag{6}$$

where $(p_t, p_k) \sim CS$ denotes the pair is sampled from the cross-style generation stream.

Here, we would like to make two remarks. First, since $C$ takes the latent feature from $F_f$ as input, we avoid the non-differentiability of the text decoding mechanism and can directly train $F_f$. Second, despite that $C$ is pretrained using feature extracted from $H_f$, we use the feature extracted from $F_f$ as input. We can perform this operation not only because these two features have the same dimension but also because we enforce them to have a similar distribution via optimizing the GAN loss, discussed below.

**GAN loss** $L_{GAN}$ is used to match the distributions of the features generated by $F_f$ and those generated by $H_f$, respectively, as processing the cross-style generation stream. We use a latent GAN formulation where we train a GAN discriminator $D$ to differentiate features extracted from $F_f$ to $H_f$. The GAN loss is given by

$$E_{p_t, p_k} \left[ - \log \left( D(H_f(p_k)) \right) - \log \left( 1 - D(F_f(f(p_t), p_k)) \right) \right]. \tag{7}$$

We realize the discriminator $D$ using a -based transformer network.

### 4.3 Generator Design

We realize the style encoder $F_s$ using a GPT-2-based transformer identical to $H_f$. After extracting a representation $z_t$ for each token $t$ in $s$, we utilize a 3-layer position-wise fully-connected network to obtain the final style code $z$ as illustrated in Fig. 2. The text decoder $F_g$ is also a GPT-2-based transformer identical to $H$. We initialize the weights in $F_s$ and $F_g$ using the weights in the pretrained $H$. Next, we compare four different ways of injecting outputs from $F_s$ into $F_g$, which represent different inductive biases and result in difference performances.

**Model A: style code as a bias to the input.** In this model, the style code $z = F_s(s)$ is directly summed up with the token-embedding and position embedding before inputting to the first transformer module in $F_g$. In other words, the input to the first transformer module in $F_g$ is $e^w_r + e^p_r + z$ where $e^w_r$ denotes as the $t$th word embedding, and $e^p_r$ denotes as the $t$th position embedding.

**Model B: style code as a summarization token.** In this model, the computed style code $z = F_s(s)$ is treated as a special token that is inserted to the beginning of the input sequence and is directed fed in the first transformer module in $F_g$. That is the input sequence length becomes $T + 1$. This design is motivated by the traditional sequence-to-sequence modeling techniques (Chung et al. (2014); Cho et al. (2014); Sutskever et al. (2014); Bahdanau et al. (2016); Vinyals & Le (2015)).
Model C: style-aware self-attention. In this model, we input \( z \) into each self-attention layer in \( F_g \) to influence its computation given by 
\[
\text{Softmax}(q_m^T k_{m-1} \sqrt{B}) v_{m-1} = \eta_m(z)
\]
where \( \eta_m \) denotes an affine transformation, \( k_{m-1} \) and \( v_{m-1} \) denotes the key and value embeddings from the \((m-1)\)th hidden layer, and \( B \) denotes the hidden dimension.

Model D: adaptive layer normalization. Inspired by the recent success in image generation tasks (Park et al. (2019); Karras et al. (2019)), we utilize the style code to modulate the hidden representations within the text decoder via normalization layers. Specifically, we replace the scale and bias parameters in the affine transformation step of the layer normalization (Ba et al. (2016)) with a style code determined scale and bias. That is
\[
\gamma^a_{m,c,t} h_{m,c,t}^a - \mu^a_{m,t} + \beta^a_{m,c,t}(z),
\]
where \( h_{m,c,t}^a \) denotes the \( c \)th hidden representation of the \( t \)th token at the \( m \)th transformer layer. We note \( a = \{1, 2\} \) since there are two layer normalization layers in each transformer in our implementation. The mean and deviation \( \mu_{m,t}^a \) and \( \sigma_{m,t}^a \) are computed across the channel dimension.

We illustrate how these models inject \( z \) to \( F_g \) in Fig. 2. In Section 5 we compare the performance of these variants and show that Model D achieves the best style generation performance.

5 Experiments

Implementation. We set the latent dimension \( B \) to 768, number of attention-heads to 16, number of transformer layers \( M \) to 16, number of tokens in a paragraph \( T \) to 512, and the vocabulary size \( V \) to 50,257 using BPE-encoding (Sennrich et al. (2015)) vocabulary from Radford et al. (2019) throughout all the models and experiments. We use a pretrained GPT-2 model \( H \) and a style comparator \( C \) in our framework. The training details of these two models are given in Appendix B.

Datasets. We compare competing methods using two newly composed datasets based on (Zhu et al. (2015); Zellers et al. (2019); Santiago (2015); See et al. (2017)).

3-Style. The dataset consists of documents from the RealNews dataset (Zellers et al. (2019)), the BookCorpus dataset (Zhu et al. (2015)), and the Reviews dataset (Yelp (2019); McAuley &
style control, it has a poor style score. On the other hand, S-GPT-2 style scores for the 3-Style G-GPT-2 the best. As expected, more superior the model is. From the figure, we found that among our models, Model D performs baselines on the 3 styles, including Sciences, Sport, Politics, Business, Technology, Entertainment, Opinion, Life, and News. Then, we divide the documents in BookCorpus into 8 different styles, which are Romance, Fantasy, Sciencefiction, Childrensbooks, Thriller, Adventure, Poetry, and Plays. We split the documents into multiple small documents by extracting the dialogues except for the Poetry and Plays. We divide the Review dataset into 3 styles, namely Yelp, Hotel, and Movie. Finally, we crawl 0.77M lyrics from http://www.azlyrics.com/. The total dataset has 35.5M documents. We hold out 3.55M documents as the validation set and 21K documents as the testing set.

Auto-evaluation metrics. We evaluate different models using fluency score, style score, style diversity score, and content novelty score. The fluency score measures whether the output paragraph reads like a human-written one. The style score checks whether the output text carries the target style. Our framework supports multimodal outputs (Huang et al. 2018). For the same input context but different reference examples of the same style, our framework should produce different output texts but all with the same style. To measure how different these outputs are, we use the style diversity score. Finally, the content novelty score is used to measure the difference between the output and the reference example. A model that directly duplicates the reference to the output is undesirable. The details of these automatic evaluation metrics are available in Appendix C.

Human study settings. We use the Amazon Mechanical Turk (AMT) platform for user studies. We conduct two studies where one evaluates fluency of the generated paragraphs while the other verifies the style correctness. For the fluency study, we present a human-written text and a machine-generated text in random order and ask the worker to choose which one is written by a human. For this metric, the closer the preference score to 50%, the better the performance.

For the style study, we perform two tests. In one test, we present a worker a generated paragraph that supposes to be in the target style. We also give the worker two human-written reference paragraphs where one is with the target style while the other is not. We then ask the worker to choose which reference paragraph has a style more similar to the generated one. In the other test, we again present a worker a generated paragraph but this time with the style categorical labels to choose from instead of the reference paragraphs. We compute the frequency that the worker selects the right style. The higher the score, the better the performance. More details are in Appendix D.

Strong baselines. We compare our framework to three strong baselines, namely the general GPT-2 model (G-GPT-2), a baseline consists of multiple style-specialized GPT-2 models (S-GPT-2), and the style-code encoding (SC) method based on the description in Keskar et al. (2019). G-GPT-2 is trained on the entire dataset using $L_{CM}$. It does not allow style control but can generate fluent texts. In S-GPT-2, we train a GPT-2 model per style. As training a GPT-2 model is costly, we only use this baseline for the 3-Style dataset evaluation. In SC, an one-hot encoding of the style class label is used as a special token for style-controllable paragraph generation. Unlike the proposed method that extracts the style code from the input paragraph, SC input the style label. The rest of the model is similar to our Model B without the style encoder.

5.1 Results

In Fig. 3, we plot the fluency and style scores achieved by our models as well as those by the baselines on the 3-Style and 21-Style datasets. The closer the model to the top-right corner, the more superior the model is. From the figure, we found that among our models, Model D performs the best. As expected, G-GPT-2 achieves the best fluency score. However, since it does not support style control, it has a poor style score. On the other hand, S-GPT-2 achieves good fluency and style scores for the 3-Style dataset. This is understandable as it utilizes a GPT-2 model for each style. However, such an approach does not scale well as GPT-2 training is expensive. We also found that SC does not achieve good style score and is inferior to our models. We suspect this is because the one-hot style class code is largely ignored during inference. Since Model D performs the best
in our framework, for the rest of the paper, we use it as our representative model for performance comparison as well as ablation study.

In Tab. [1] we show several generation results from our Model D. We find that the output texts are fluent and respect the styles of the references. More output examples are available in Appendix [E].

In Tab. [2] we show the style diversity scores achieved by our models. We found that all of our 4 models can generate diverse styled paragraphs conditioning on the same context and different reference examples with the same style.

**Human evaluation.** In Tab. [3] we report user study results on fluency and style control. We found that our model achieves great fluency on both of the datasets. Compared to SC, our model performs better in controlling the style in the output texts.

**Ablation study.** We conduct an ablation study on the loss terms in the proposed objective function and report the results in Tab. [4] using the 21-Style dataset. The results show that each term is important. Removing $L_{DIST}$ leads to a degraded content novelty score. Removing $L_{STYLE}$ leads to a degraded style score, though an improved fluency score and a content novelty score. Removing $L_{GAN}$ leads to both degraded fluency and style diversity scores.

### 6 CONCLUSION

We presented a language generative framework for style example-guided paragraph generation. To the best of our knowledge, we were the first to achieve such style-controllability on paragraph generation. We attributed the success to our carefully designed learning objective function, the generator network, and the newly composed large-scale dataset consisting of documents of various text styles.
REFERENCES


Under review as a conference paper at ICLR 2020


Andrew L Maas, Raymond E Daly, Peter T Pham, Dan Huang, Andrew Y Ng, and Christopher Potts. Learning word vectors for sentiment analysis. In *Annual Meeting of the Association for Computational Linguistics (ACL)*, 2011.


Sudha Rao and Joel Tetreault. Dear sir or madam, may I introduce the yafc corpus: Corpus, benchmarks and metrics for formality style transfer. In *Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2018.


A Hyper-parameters Tuning

We tune the hyper-parameters on a pilot-test dataset. This pilot-test dataset has 3K training examples and 400 hold-out examples. We perform a grid search on log-scale. We utilize the Frechet Embedding Distance (Semeniuta et al., 2018) to select best hyper-parameters. For \( \lambda_{DIST} \), \( \lambda_{STYLE} \), and \( \lambda_{GAN} \), we try \{1.0, 0.1, 0.1\}, \{0.1, 0.1, 0.1\}, and \{1.0, 0.01, 0.01\}. For betas used in Adam Optimizer, we try \{0.9, 0.999\} and \{0.0, 0.9\}. For weight decay, we try 0.01 and 0.0. For the initial learning rate, we try 0.00025 and 0.000025. Eventually, we use 0.00025 as initial learning rate and \{1.0, 0.01, 0.01\} for \( \lambda_{DIST} \), \( \lambda_{STYLE} \), as well as \( \lambda_{GAN} \) for all models except for the Model B. We use \{1.0, 0.01, 0.1\} for the Model B instead.

B Pretraining of H and C

Pretraining of the GPT-2 model H. We pretrain H on our collected dataset \( \mathbb{D} \) from scratch. We use Adam optimizer with 0.00015 initial learning rate, \( \beta_1 \) and \( \beta_2 \) are set by \{0.9, 0.999\}, cosine learning rate decay style, and 0.01 weight decay. The batch size is set to 512. The total training iterations is 320K which takes 2 weeks.

Pretraining of the style comparator C. We pretrain the Style Comparator C using 256 and 512 batch size, respectively. The initial learning rate is 0.00015 with 0.01 weight decay and cosine learning rate decay style. The optimizer is also Adam with \( \beta_1 = 0.9 \) and \( \beta_2 = 0.999 \). Since the training converges very quickly, we do early stopping if the accuracy on validation set plateaus. Eventually, we use the checkpoint at 28K and 50K iterations to train on the 3-style and 21-style datasets, respectively. The accuracy on hold-out validation set are 98.8% and 87.8% on 3-styles and 21-styles datasets, respectively.

C Auto-evaluation Metrics

Fluency score. To ensure the generated paragraph is fluent and coherent, we utilize a pretrained GPT-2 model to measure the perplexity of the generated paragraph. We compute the generation likelihood over each token using the model and treat the generated paragraph \( y_{0:T-1} \) as inputs and \( y_{1:T} \) as labels. Given the input text, the pretrained GPT-2 returns the probability distribution of next token over the vocabulary. Then, we measure the perplexity by this probability distribution and label. Since our dictionary size is 50257, the random guess of the next token would result in 50257 perplexity (\( \mathbb{L}_{LM} = \log(50257) \approx 10.83 \)). Thus, we set 10.83 as an upper bound and define the fluency score of the generated paragraph as \( 10.83 - \log(\text{perplexity}) \). In this sense, a higher fluency score means lower perplexity.

Style score. We train 3/21 binary style classifiers (since we have 3/21 different styles in the 3-Style/21-Style dataset) by finetuning a GPT-2 network to automatically evaluate whether the generated text carries the style of a target class. These 3/21 classifiers achieve average classification accuracies of 99.1%/96.3%. During the testing phase, for a target style, if the corresponding style classifier correctly predicts 1 for the generated paragraph computed by a model, we count it as a successful trial. We compute the success rate over the test set and use the result as the style score for the model.

Style diversity score. We adopt the LPIPS distance (Zhang et al., 2018) to measure the diversity of the generation outputs conditioning on the same context. To implement this metric, we first extract the feature representation from each token in a generated paragraph by a pretrained GPT-2 model. We compute the mean representation of the tokens in a paragraph as the paragraph-level representation. Then, we measure the \( L_2 \) distance between two paragraph-level representations of two different paragraphs generated using the same context but two different references written in the same style. In this sense, a larger distance value implies the styles of the two generated paragraphs are more different.

To get an idea of the range of this metric, we compute an upper bound and a lower bound. We consider two paragraphs from two documents of different styles should have a high style diversity score. We hence sample 1000 paragraphs from each style and use the pretrained GPT-2 model to extract deep features. After taking average over the token-dimension, we obtain 1000 by 768
representation for each style. Then, we compute the $L_2$ distance between of these matrices divided by 1000. This gives us a matrix of size $21 \times 21$ measuring the pairwise distance between two styles. We use the largest value in this matrix as our upper bound, which is 15.67.

For the lower bound, since two different paragraphs from the same document should have a low style diversity score, we use their scores for the lower bound computation. Specifically, we compute the average distance between two different paragraphs from the same document. We do this for each style and obtain 21 different values. We obtain the lower bound by taking average over these values, which is 4.52.

**Content novelty score.** To verify that our model is not simply duplicating the content from reference paragraph, we utilize LPIPS distance (Zhang et al. (2018)) to measure the difference between the generated paragraph and the input reference paragraph. We again use a pretrained GPT-2 model for extracting a feature representation for each token. To compute the distance between two paragraphs, we compute the bipartite matching cost between the tokens from the two paragraphs. Specifically, we first compute the $L_2$ distances between any token representation in one paragraph to all the token representations in the other paragraph. We then compute the minimum cost assignment by solving a bipartite matching problem. In order to get an idea about the range of the content novelty score, we compute an upper bound and a lower bound using a similar approach as the one used for the style diversity score. We find the upper bound value is 31.46 and the lower bound value is 15.13.

**D Human Evaluation**

To participate in our user study, a worker has to be awarded the Master Qualification by AMT and has at least 97% life-long HIT approval rate. We generate 2 paragraphs for a context with 2 different reference paragraphs from the same style. For 3-style dataset, we randomly sample 50 examples from each style and do style generation for all 3 styles. Thus, there are 900 examples for each testing model. For 21-style dataset, we randomly sample 42 examples from testing set and do style generation on all 21 styles. Thus, there are 1764 examples for each testing model. There is a typical example for experiment of fluency human study in Fig. 4.

![Figure 4: A typical example of the questionnaire for our A/B test on AMT for the fluency study.](image)

**E Additional Results**

In Tab. 5, 6, 7, 8, and 9, we present additional example outputs of the proposed framework. The results demonstrate that the proposed model is fully capable of generating fluent output texts respecting the style of the input reference examples.
### Context – Poetry Style

If anything, I’m a petty thief in a world of forgetting, a blade in an elm, a bottled note in the sea of Victory Boulevard at dusk. Friends, family, gallows saints & ghosts, rig a swing at the edge of the shore when I’m gone for the birds to perch & the bats to glide through, to disappear & appear in the small places the moon cannot touch. Thank you for the amorous & discontented beauty you saved for me here, though most days, my silence was a soldier’s slogged hand hiding a match. I love you. I leave you the match.

### Reference – Lyrics Style (Senorita)

I love it when you call me senorita
I wish I could pretend I didn’t need ya
But every touch is ooh la la la
It’s true, la la la
Ooh, I should be running
Ooh, you keep me coming for you
Land in Miami
The air was hot from summer rain
Sweat dripping off me
Before I even knew her name, la la la
It felt like ooh la la la

### Generated

I turn your body in my lap
your purse, a Pen, of a beautiful people
and I love you the way you love me, look at you drop a hand

You wake up your ad libs Cirospis, cetachi, cetachi, cetachi, cetachi, massageas, cetachi, cetachi, massageas, London Cirospis
I love you inside, I love you at the foot of your bed,
I love you inside, I love you at the foot of your bed,

Leave your big collar, you leave your small collar, you leave your little collar

Your life is just a half hour, it’s your dance courses
you take a high vase of [real] ciche
just like our flowers

We all want it so much, we want it so much, everybody
Then kiss thing, girl go kiss it

I set the girl on her arms
against the state where no one could give it
–love is reserved for us, romanceis reserved for us, romanecis reserved for us, lovesis reserved for us, lovesis reserved for us,

And it began at the state where no one could give it
–love is reserved for us, romanceis reserved for us
–you are not just dresses but you are wearing them all a million times
, you are your own beautiful girls’ effects
They blend in faster than the others
You meet and love all things, you . . .

I am a proud Greek, an ancient Greek virgin
And I, like you, in high places, love all things, sexy
But on your early morning sleep, samba but in high places, sexy
Vešey y Vešey feather savant

### Context – News Style

CARL AZUZ, CNN 10 ANCHOR: Hi, I’m Carl Azuz and you’re watching CNN 10. Ten minutes of news explained. There are a lot of fast moving political headlines from the United States right now. On Wednesday the White House is expected to release the transcript of a phone call. A conversation between the presidents of the United States and Ukraine. The circumstances surrounding that call and something allegedly discussed on it were the reasons why on Tuesday Democrats from the U.S. House of Representatives announced they were moving forward on an impeachment inquiry, an investigation.

### Reference – Lyrics Style (Baby Shark)

Baby Shark doo doo, doo doo doo doo
Baby Shark doo doo, doo doo doo doo
Baby Shark doo doo, doo doo doo doo
Baby Shark doo doo, doo doo doo doo
Baby Shark doo doo, doo doo doo doo
Baby Shark doo doo, doo doo doo doo

### Generated

of a real, if not hoaxed, race in the United States.
All right, hey can I have a snack, Bob Bush?
A little snack baby!

Measures how to speed up the Country children?
Programs be grateful, Goodwill be giving.
Messages like, what is this thing about 8 am in the Middle?
Benjamin Franklin? Go
Nintendo Wii, Mario
Nintendo DS, U
Maria Barbara, Nice Guy
Johnny Cash, Mickey Mouse 2, Magic Tony 3, Superman, Backzapper, Yellowjack, Bambi, Pig
His sister Diva, Saaath, Lil Saaath
A sad, sad, sad
If we only do than, what are he doing with it ?
Let’s move forward and make it better, we can officially
make it good good, especially now Delma 0, Cadillac .
MAHA YEAH!!!!! :DDD
La Dara No Art, Chicken and Beach
(#Creats La Dara No Art, Chicken and Beach)

Table 5: Qualitative Results: to lyrics style.
Table 6: Qualitative Results: to reviews style.
Context – Children’s Books Style (Snow White and the Seven Dwarfs)
Once upon a time . . . in a great castle, a Prince’s daughter grew up happy and contented, in spite of a jealous stepmother. She was very pretty, with blue eyes and long black hair. Her skin was delicate and fair, and so she was called Snow White. Everyone was quite sure she would become very beautiful. Though her stepmother was a wicked woman, she too was very beautiful, and the magic mirror told her . . .

Reference – Poetry Style
A black vulture crushes my bones.
A white vulture in a photograph.
Houses open their mouths, swallow flowers. Lightning doesn’t choose. I brushed the spruces by hand. A car has a tire, a farmland a whip, trash cans are falling. We have two natures. The vulture circles the belly at Pujiaka Gora. Gingerbread hearts, red blood, something awaits! Down bangs into the bricks of Brooklyn. I’m still asleep in silvery milk. When the wind strikes the grass, I’ll be in front of the house. Crickets will burn the air. The sound will be like in the hall where there’s plush lining in Christ’s tomb.

Generated
wait for my goddaughter to faint! For my stepmother is the thought of greenery, gardens and owyls!
Still having a conversation scraped off the edge of my own couch,
I hear the Ming river cry over the hearth,
A tide will strawn,
and so will flowers.
Up me through her hair,
More flowers,
A thistle bleeds,
a unite of vibrant sugar.
A red half of my lungs open and reach my face,
Several holes in my chest.
And a piano

Context – Instagram Style
An amazing meal had at @thestratford, testing their newly opened restaurant. Buttermilk deep-fried chicken with a doughnut and salad. An absolute must. Thank you @thestradford #thestradford #londonrestaurant #londonrestaurants #cuisine #food #foodporn #foodphotography #foody #foodgasm #foodstagram #foodlover

Reference – Poetry Style
A black vulture crushes my bones.
A white vulture in a photograph.
Houses open their mouths, swallow flowers. Lightning doesn’t choose. I brushed the spruces by hand. A car has a tire, a farmland a whip, trash cans are falling. We have two natures. The vulture circles the belly at Pujiaka Gora. Gingerbread hearts, red blood, something awaits! Down bangs into the bricks of Brooklyn. I’m still asleep in silvery milk. When the wind strikes the grass, I’ll be in front of the house. Crickets will burn the air. The sound will be like in the hall where there’s plush lining in Christ’s tomb.

Generated
shows in the building news feeds, adorned with neon signs and mason jars. Just how much of the thinkers are grating compassion, you know thoughts, for weeks burning in the hemispheres. I call my guests ‘the french’ s, the eggs and cream, the eggs and chicken
The Kinrear growling embers of bees tingle inside each woman. ”Brat a cusa a crow.”
She pauses, looking at Anna, doubt sitting heavy on her face. “You are warm as a fly. For you, those who want to live you will be happy, .

Context – Sport News Style
Stating the obvious apparently costs $50,000 in the NBA.
According to Shams Charania of Stadium and The Athletic, the league fined the Milwaukee Bucks after general manager Jon Horst told fans at a town hall event the Bucks plan to offer Giannis Antetokounmpo a supermax extension when they’re eligible to do so.
Horst referenced how league rules prevent Milwaukee from formally discussing specifics regarding what it would .

Reference – Poetry Style
It is written somewhere in the sky we are supposed to forget quickly. How? An elephant stands on a tiny pedestal. In a field under a tent. The strongman spectacular in his suit and topcoat, expert in redirection, points towards the acrobat. The acrobat astounds because he could fall.
The other day I told my friend the story in which you fell through the floorboards of an old barn. As I told the story, I recalled the manner in which you carefully transferred milk to a newborn calf. You beckoned to Babette singing come Rosy Boss come Rosy.

Generated
earmuged its way in May far a warrant to call.
After the conditions were waived, Indiana Pacers head coach Luke Walton and Bucks forward Patrick Ewing came into the meeting to see what those rules meant.
So what were the rules?
At first, we thought between letting training and verbal communication come into play. Now, with the way you communicate, you just have to make some other national request.
There’s a tall castle rising in the sky, and a great sweet tale conjured up right. I cover a chandelier in a gift box; over the ledge you can spot a diamond with feathers on it. Learn to try again.
We can’t afford to

Table 7: Qualitative Results: to poetry style.
Under review as a conference paper at ICLR 2020

<table>
<thead>
<tr>
<th>Context – Random sentence</th>
<th>Reference – Children Books Style (Little Red Cap)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have a cute dog</td>
<td>&quot;Good day to you, Little Red Cap.&quot;</td>
</tr>
<tr>
<td></td>
<td>&quot;Thank you, wolf.&quot;</td>
</tr>
<tr>
<td></td>
<td>&quot;Where are you going so early, Little Red Cap?&quot;</td>
</tr>
<tr>
<td></td>
<td>&quot;To grandmother’s.&quot;</td>
</tr>
<tr>
<td></td>
<td>&quot;And what are you carrying under your apron?&quot;</td>
</tr>
<tr>
<td></td>
<td>&quot;Grandmother is sick and weak, and I am taking her some cake and wine. We baked yesterday, and they should be good for her and give her strength.&quot;</td>
</tr>
<tr>
<td></td>
<td>&quot;Little Red Cap, just where does your grandmother live?&quot;</td>
</tr>
<tr>
<td></td>
<td>&quot;Her house is good quarter hour from here in the woods, under the three large oak trees. There’s a hedge of hazel bushes there. You must know the place;&quot; said Little Red Cap.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Generated</th>
<th>Reference – Children Books Style (Snow White and the Seven Dwarfs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Well, I found a small white dog worth the trouble.&quot;</td>
<td>Once upon a time . . . in a great castle, a Prince’s daughter grew up happy and contented, in spite of a jealous stepmother. She was very pretty, with blue eyes and long black hair. Her skin was delicate and fair, and so she was called Snow White. Everyone was quite sure she would become very beautiful. Though her stepmother was a wicked woman, she too was very beautiful, and the magic mirror told her this every day, whenever she asked it.</td>
</tr>
<tr>
<td>&quot;What species are my family?&quot;</td>
<td>&quot;Mirror, mirror on the wall, who is the loveliest lady in the land?&quot; The reply was always &quot;You, your Majesty,&quot; until the dreadful day when she heard it say, &quot;Snow White is the loveliest in the land!&quot; The stepmother was furious and, wild with jealousy, began plotting to get rid of her rival. Calling one of her trusty servants, she bribed him with a rich reward to take Snow White into the forest, far away from the Castle. Then, unseen, he was to put her to death. The greedy servant, attracted to the reward, agreed to do this deed, and he led the innocent little girl away. However, when they came to the fatal spot, the man’s courage failed him and, leaving Snow White sitting beside a tree, he mumbled an excuse and ran off. Snow White was all alone in the forest. Night came, but the servant did not return. Snow White, alone in the dark forest, began to cry bitterly. She thought she could feel terrible eyes spying on her, and she heard strange sounds and rustlings that made her heart thump. At last, overcome by tiredness, she fell asleep curled under a tree. Snow White slept fitfully, waking from time to time with a start and staring into the darkness round her. Several times, she thought she felt something, or somebody touch her as she slept. At last, dawn woke the forest to the song of the birds, and Snow White too, awoke. A whole world was stirring to life and the little girl was glad to see how chilly her fears had been. However, the thick trees were like a wall round her, and as she tried to find out where she was, she came upon a path. She walked along it, hopefully. On the walked till she came to a clearing. There stood a strange cottage, with a tiny door, tiny windows and a tiny chimney pot. Everything about the cottage was much larger than it ought to be. Snow White pushed the door open.</td>
</tr>
<tr>
<td>&quot;I suppose ten,&quot; he said</td>
<td>&quot;Mirror, mirror on the wall, who is the loveliest lady in the land?&quot; The reply was always &quot;You, your Majesty,&quot; until the dreadful day when she heard it say, &quot;Snow White is the loveliest in the land!&quot; The stepmother was furious and, wild with jealousy, began plotting to get rid of her rival. Calling one of her trusty servants, she bribed him with a rich reward to take Snow White into the forest, far away from the Castle. Then, unseen, he was to put her to death. The greedy servant, attracted to the reward, agreed to do this deed, and he led the innocent little girl away. However, when they came to the fatal spot, the man’s courage failed him and, leaving Snow White sitting beside a tree, he mumbled an excuse and ran off. Snow White was all alone in the forest. Night came, but the servant did not return. Snow White, alone in the dark forest, began to cry bitterly. She thought she could feel terrible eyes spying on her, and she heard strange sounds and rustlings that made her heart thump. At last, overcome by tiredness, she fell asleep curled under a tree. Snow White slept fitfully, waking from time to time with a start and staring into the darkness round her. Several times, she thought she felt something, or somebody touch her as she slept. At last, dawn woke the forest to the song of the birds, and Snow White too, awoke. A whole world was stirring to life and the little girl was glad to see how chilly her fears had been. However, the thick trees were like a wall round her, and as she tried to find out where she was, she came upon a path. She walked along it, hopefully. On the walked till she came to a clearing. There stood a strange cottage, with a tiny door, tiny windows and a tiny chimney pot. Everything about the cottage was much larger than it ought to be. Snow White pushed the door open.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Context – Yelp Style</th>
<th>Reference – Children Books Style (Snow White and the Seven Dwarfs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Someone is easily one of my go-to spots for dessert because there’s no lactose in the ice cream, it’s very Instagrammable, and they have rotating flavors.</td>
<td>Once upon a time . . . in a great castle, a Prince’s daughter grew up happy and contented, in spite of a jealous stepmother. She was very pretty, with blue eyes and long black hair. Her skin was delicate and fair, and so she was called Snow White. Everyone was quite sure she would become very beautiful. Though her stepmother was a wicked woman, she too was very beautiful, and the magic mirror told her this every day, whenever she asked it.</td>
</tr>
<tr>
<td>My friend and I decided to visit this location a few days after it opened and surprisingly, there was a really short line. This location had 6 flavors (matcha, milk, oreo, ube, coffee, and milk tea). I tried the matcha flavor and it was delicious!</td>
<td>&quot;Mirror, mirror on the wall, who is the loveliest lady in the land?&quot; The reply was always &quot;You, your Majesty,&quot; until the dreadful day when she heard it say, &quot;Snow White is the loveliest in the land!&quot; The stepmother was furious and, wild with jealousy, began plotting to get rid of her rival. Calling one of her trusty servants, she bribed him with a rich reward to take Snow White into the forest, far away from the Castle. Then, unseen, he was to put her to death. The greedy servant, attracted to the reward, agreed to do this deed, and he led the innocent little girl away. However, when they came to the fatal spot, the man’s courage failed him and, leaving Snow White sitting beside a tree, he mumbled an excuse and ran off. Snow White was all alone in the forest. Night came, but the servant did not return. Snow White, alone in the dark forest, began to cry bitterly. She thought she could feel terrible eyes spying on her, and she heard strange sounds and rustlings that made her heart thump. At last, overcome by tiredness, she fell asleep curled under a tree. Snow White slept fitfully, waking from time to time with a start and staring into the darkness round her. Several times, she thought she felt something, or somebody touch her as she slept. At last, dawn woke the forest to the song of the birds, and Snow White too, awoke. A whole world was stirring to life and the little girl was glad to see how chilly her fears had been. However, the thick trees were like a wall round her, and as she tried to find out where she was, she came upon a path. She walked along it, hopefully. On the walked till she came to a clearing. There stood a strange cottage, with a tiny door, tiny windows and a tiny chimney pot. Everything about the cottage was much larger than it ought to be. Snow White pushed the door open.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Generated</th>
<th>Reference – Children Books Style (Snow White and the Seven Dwarfs)</th>
</tr>
</thead>
</table>
| Designs concerned: Some hail the hero, Verdicts seem to be the wrath of the storm. The main reason sweets can’t provide a lot of herace; One side calls her errible, the other side tells her, "Watch out, this is a great purity factory.""Oh, she we are degraded," a low voice whispered into her, "the delight heights thus are deeds rather than love.""For only things that are great, and that happiness cannot wast; For she appears so what is best. In this changing season, Regan might win a daily battle, In the surprise that revealing her majesty brings. At the end of the day, surprised and amused and bewildered," says Regan, "What did I wait for," demands a woman, "burst upon your beautiful mother!" The cease goes on and on as they argue over the choices best, Even though the simpler justice began to prevail, the evil abyss began to mend. "Such is her heart she didn't even know that, The mercy of her master she feared would take her life," she said, "a woman who looked only for herself. She pretended to think the choice that befell him was her love;But true, she turned to the figure in her bed."
<table>
<thead>
<tr>
<th>References</th>
<th>Reference – Children Books Style (Snow White and the Seven Dwarfs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweetest song, on the one side of which is the inner figure in the solution to her woes; the sensual dictates; by which she might follow the recipe of the book of grateful dreams, The sweetest song I ever had ever heard, to this new flourishes. Perfect!&quot; she said to Violet, &quot;not just found but realised it had mean</td>
<td></td>
</tr>
</tbody>
</table>

Table 8: Qualitative Results: to children books style.
Context – Movie Reviews Style
I am struggling to find a good reason to give you in order to see this movie, but I cannot. I will not stick to how the actors were playing because it was not their fault about the role they had been given. On the other hand I will base my review on the scenario. What was the meaning of the movie? Which was the target of the plot? I did not manage to answer neither of the aforementioned quizzes. I read that

Reference – Politic News Style
(CNN)President Donald Trump is wasting no time in attempting to torch House Speaker Nancy Pelosi’s impeachment gamble in a battle that will define his presidency and the 2020 election.

Trump has pledged to publish an un-redacted and declassified transcript on Wednesday of a phone call with Ukraine’s leader to reject the Supreme Court that motivated him to do what, instead of taking a stance that is less adversarial, may be demagogic.

The transcript and the congressional reaction to the whistleblower’s report could be critical in establishing the early stages of the impeachment fight and to shaping public opinion that will ultimately dictate how it turns out.

Regardless of the outcome, Trump finds himself at the center of a rare and historic showdown as only the fourth president in US history to face the realistic threat of impeachment.

Generated
a short about a World War II president contemplating a secret meeting with Hitler won’t be released, but the person being asked that it is, in fact, a statement he made in Munich on May 6, 1938 that would lead to the end of World War II.

I also assume that modern life in the present moment right now is much more uncomfortable, and it’s difficult for my migrator to provide empathy because what narratives I have are so blended with existing views that the idea of Trump president is only in the realm of his childish note to reject the Supreme Court that motivated him to do what, instead of taking a stance that is less adversarial, may be demagogic.

Trump by comparison isn’t such a set up to gain the desired political stature that Monty Python’s sketch computer shows him. Nor, it’s just a new territory he’s taking to be killed. Contentious and creative questions are looming. The real question is whether all this will ultimately lead to the abolition of the Supreme Court.

Context – Paper Style (Abstract of this paper)
We introduce a language generative model framework for generating a styled paragraph based on a context sentence and a style reference example. The framework consists of a style encoder and a text decoder. The style encoder extracts a style code from the reference example, and the text decoder generates texts based on the style code and the context. We propose a novel objective function to train

Reference – News Style
(CNN)Massachusetts Gov. Charlie Baker has called for a temporary statewide ban on the sale of all e-cigarettes and vaping products in response to a nationwide outbreak of lung injuries associated with vaping.

"I’m officially declaring a public health emergency in the Commonwealth due to severe lung disease associated with the use of e-cigarettes and marijuana-influenced vaping products," Baker said during a press conference on Tuesday.

“I’m requesting that the public health council order a four-month temporary ban on the sale of vaping products in retail establishments, online and through any other means, effectively immediately," he said. “We as a Commonwealth need to pause sales in order for our medical experts to collect more information about what is driving these life-threatening vaping-related illnesses.”

Generated
everyone through a useful tool or method; each custom bound to require instruction from an expert inside the business community. This type of exercise should be done in a timely manner and without fanfare and coming from the sports, entertainment and educational communities, and in a highly professional scientist setting, where everyone could help us understand things we need to improve want people to be able to start the line by entering into an active conversation with scientists – which is a "great idea," if, by definition, is humans for.

The medical community has an primary role in the global health and well-being of the globe beyond infections, obesity, heart disease, diabetes, and opioid addiction.

"I’m officially declaring a public health emergency in the Commonwealth due to severe lung disease associated with the use of e-cigarettes and marijuana-influenced vaping products," Baker said during a press conference on Tuesday.

"I’m requesting that the public health council order a four-month temporary ban on the sale of vaping products in retail establishments, online and through any other means, effectively immediately," he said. "We as a Commonwealth need to pause sales in order for our medical experts to collect more information about what is driving these life-threatening vaping-related illnesses.”

Context – Children Books Style (Little Three Pigs)
Once upon a time there was an old mother pig who had three little pigs and not enough food to feed them. So when they were old enough, she sent them out into the world to seek their fortunes.

The first little pig was very lazy. He didn’t want to work at all and he built his house out of straw. The second little pig worked a little bit harder but he was somewhat lazy too and he built his house out of sticks. Then, they sang and danced and played together the rest of the day.

The third little pig worked hard all day and built a new house, away from that stage, away from the theater, away from the movie.

But the fourth little pig worked hard it had to be him.

Steven Spielberg is one of us white people, whom we think of as “the lovable Dear Dad”.

In the end, Spielberg created the “Ralph” sequence with Donkey Kong.

Everyone who adopts a culture of old blood, that’s Abraham Lincoln for those of you white people.

I once invited a school animal to a movie so that I could ask him, ‘Maybe Jaws + Shaun of the Dead’, forever. She said, ‘Sure.’ ‘I’ll be your film, but I don’t want to be one of you gum and crackers any more.’ She’s this confused bunny.

(Click here to see a video line-up of Teri Garr’s past “Ralph” scenes and upcoming “Dreaming Tom Shadow.”)

Table 9: Qualitative Results: to News style.

20