Toward Controlled Generation of Text

Agustinus Kristiadi and Yonathan Santosa

University of Bonn

21 Dec 2017

Outline

Generative Model

- Recall Bayes' Rule:

$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)}$$

- Generative Model: modeling $P(X)$.
- Important models: GAN, VAE.
- In our work we will be based on VAE.

Autoencoder (AE)

- Neural nets that take input X and to reconstruct it, i.e. outputting $\hat{X}$

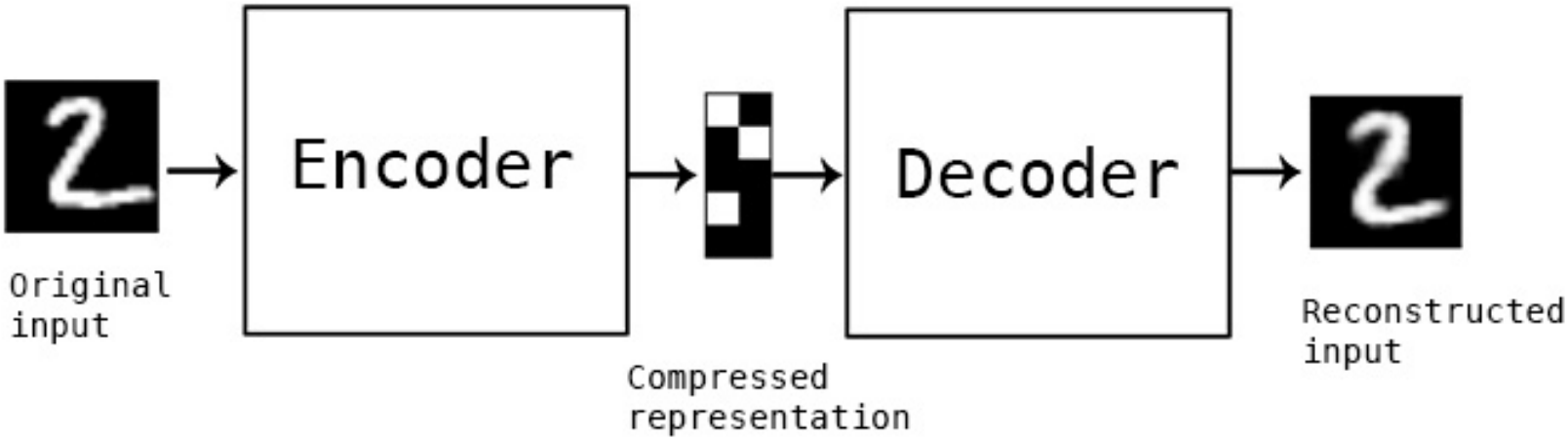

Figure 1: From: https://blog.keras.io

Variational Autoencoder (VAE) [1]

- Constraint the hidden layer into some distribution (of latent variable).
- Force the hidden layer to match our prior distribution, e.g. standard normal.
- Objective: maximize (reconstruction + hidden unit regularization)

$$\max_{\theta} \mathcal{L}(\theta; X) = \mathbb{E}_{q(z|X)}[\log p(X|z)] - D_{KL}[q(z|X) \| p(z)]$$

- In practice: $q(z|X)$ and $p(X|z)$, are neural nets.

[1] Kingma and Welling, 2013

Benefits of VAE

- Latent variables nicely contained in $\mathcal{N}(0, I)$.
- Generating data X becomes possible: $z \sim \mathcal{N}(0, I); X \sim P(X|z)$.

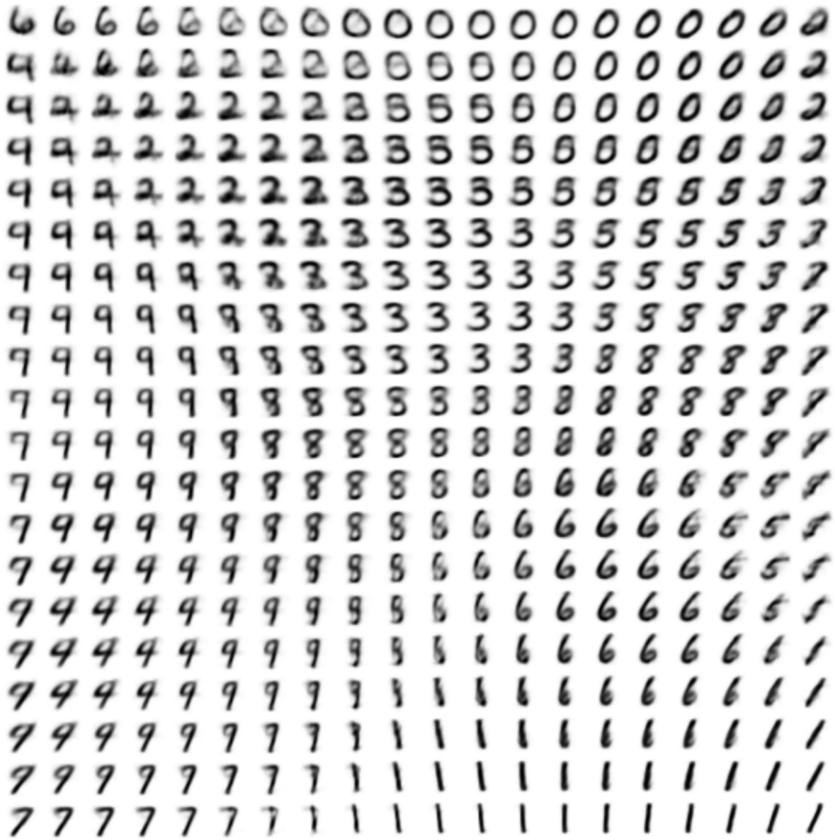

Figure 2: By interpolating z and transform it with $P(X|z)$ we can interpolate data. From: https://ermongroup.github.io/cs228-notes.

Toward Controlled Generation of Text (Hu, 2017)

- Extending VAE model
  - Use LSTM-RNNs as encoder and decoder.
  - Add another neural net to enforce conditional attribute constraint.
- Enables us to condition text generation.
  E.g. generate text with past tense and positive sentiment:
  "this was spectacular , i saw it in theaters twice".

Architecture

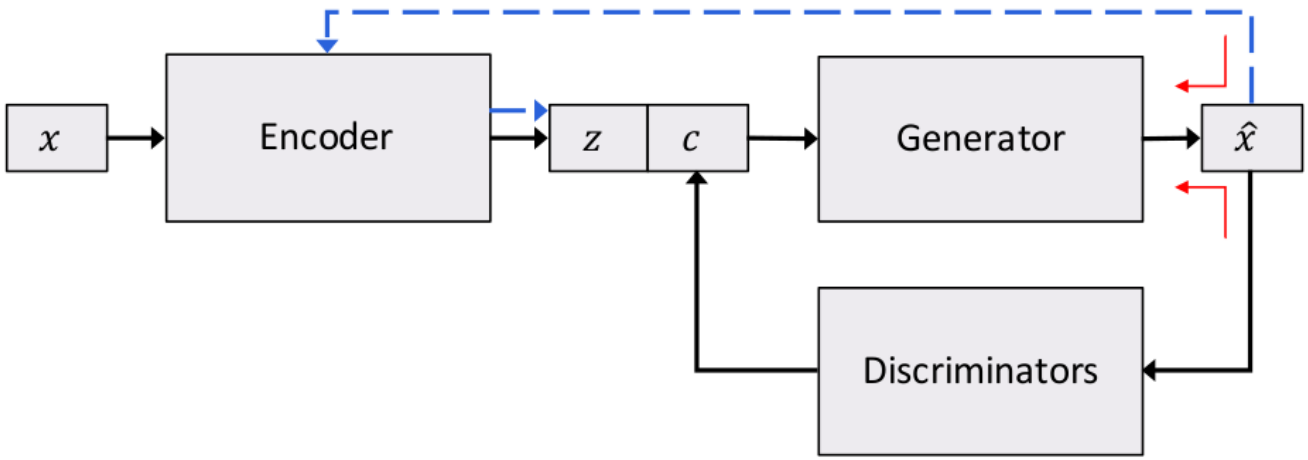

Figure 3: Hu, et al., 2017

- We are optimizing:

$$\min_{\theta_G,\theta_E} \mathcal{L}_{VAE} + \lambda \mathcal{L}_{attr}$$

$$\min_{\theta_D} \mathcal{L}_D = \mathcal{L}_s + \lambda_u \mathcal{L}_u$$

where $\mathcal{L}_{attr}$ is loss function for conditional attribute constraint.

Algorithm

Algorithm 1 Controlled Generation of Text

Input: A large corpus of unlabeled sentences $\mathcal{X} = \{\boldsymbol{x}\}$
A few sentence attribute labels $\mathcal{X}_L = \{(\boldsymbol{x}_L, \boldsymbol{c}_L)\}$
Parameters: $\lambda_c, \lambda_z, \lambda_u, \beta$ – balancing parameters

1: Initialize the base VAE by minimizing Eq.(4) on $\mathcal{X}$ with $\boldsymbol{c}$ sampled from prior $p(\boldsymbol{c})$
2: **repeat**
3: Train the discriminator D by Eq.(11)
4: Train the generator G and the encoder E by Eq.(8) and minimizing Eq.(4), respectively.
5: **until** convergence

Output: Sentence generator G conditioned on disentangled representation $(\boldsymbol{z}, \boldsymbol{c})$

Figure 4: Hu, et al., 2017

Example Expected Results

Varying the unstructured code z	
("negative", "past")	*("positive", "past")*
the acting was also kind of hit or miss .	his acting was impeccable
i wish i 'd never seen it	this was spectacular , i saw it in theaters twice
by the end i was so lost i just did n't care anymore	it was a lot of fun
("negative", "present")	*("positive", "present")*
the movie is very close to the show in plot and characters	this is one of the better dance films
the era seems impossibly distant	i 've always been a big fan of the smart dialogue .
i think by the end of the film , it has confused itself	i recommend you go see this, especially if you hurt
("negative", "future")	*("positive", "future")*
i wo n't watch the movie	i hope he 'll make more movies in the future
and that would be devastating !	i will definitely be buying this on dvd
i wo n't get into the story because there really is n't one	you will be thinking about it afterwards, i promise you

Figure 5: Hu, et al., 2017

Conclusion

- VAE is a useful modification of original autoencoder.
- We can extend VAE to also learn conditional constraint.
- We can generate text with desired properties based on the conditional constraint.

References

- Kingma, Diederik P., and Max Welling. "Auto-encoding variational bayes." arXiv preprint arXiv:1312.6114 (2013). [pdf]
- Hu, Zhiting, et al. "Toward controlled generation of text." ICML 2017. [pdf]