Overview: Adversarial Auto-Encoders for Text

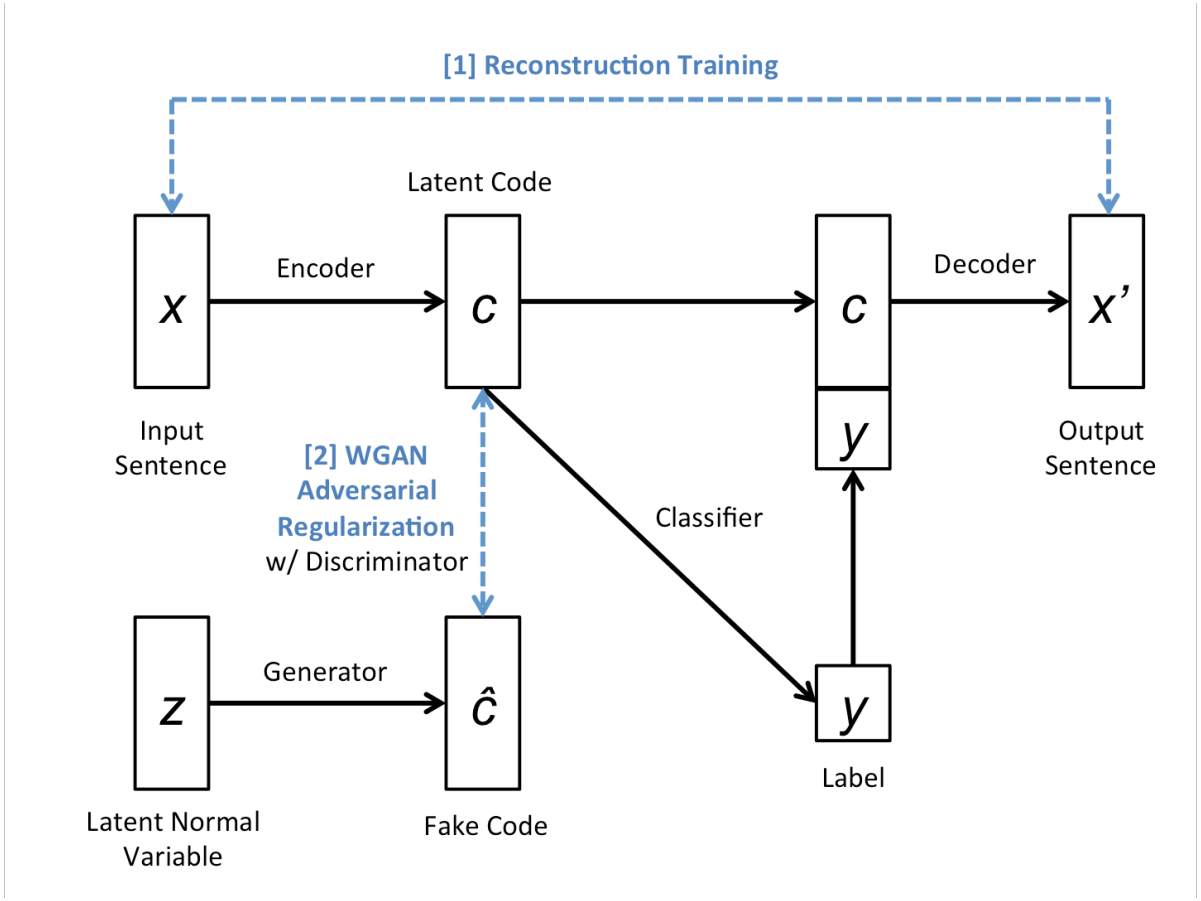
- Motivation: Learning Latent Representations for Text
  - Semi-supervised Learning
  - Text style transfer

	ORIGINAL TEXT	MODERN TEXT
	JULIET	**JULIET**
	By and by, I come.—	Alright, I'm coming!—I beg you to stop trying for me and
	To cease thy strife and leave me to my grief.	leave me to my sadness. Tomorrow I'll send the
155	Tomorrow will I send.	messenger.
	ROMEO	**ROMEO**
	So thrive my soul—	My soul depends on it—

- Several challenges
  - VAEs reduce to prior-ignoring language models
  - GANs cannot deal with discrete observations
  - AAEs experience mode-collapse & repeatedly generate same samples

Adversarially Regularized Auto-Encoder (Zhao et. al, 2017)

- ARAE = discrete AAE + WGAN + learned prior
- Improvements in semi-supervised learning
- State-of-the-art on *unaligned* text style transfer



Next Idea: Disentangled ARAE

- Separate the *label* from the *code* (e.g. style) in latent space
- Use a GAN to train the label to be categorical