

# Overview: Adversarial Auto-Encoders for Text

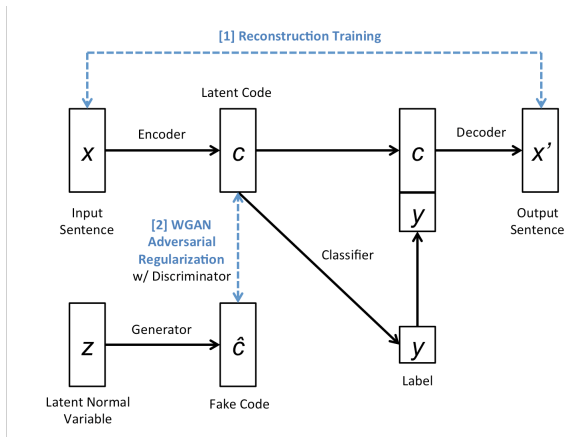
- Motivation: Learning Latent Representations for Text
  - Semi-supervised Learning
  - Text style transfer

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

- Several challenges
  - VAEs reduce to prior-ignoring language models
  - GANs cannot deal with discrete observations

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

- $ARAE = VAE + GAN$
- 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

