**Semi-supervised Learning**

SSL: Leverage limited labeled data and a large amount of unlabeled data to improve generalization.

- **Key assumptions:** Smoothness assumption, low density separation (cluster) assumption.
- **Traditional discriminative methods:** Self-training, Co-training, TSVM, SVM. Graph-based methods, regularizations such as Entropy Minimization (all explore the key assumptions)
- **Generative methods:** Gaussian Mixtures, more recently, VAEs, GANs

**Background & Motivation**

- **Main results of SNTG**
- **Semi-supervised Learning**
  - Leveraging limited labeled data and a large amount of unlabeled data to improve generalization.
  - **Key assumptions:** Smoothness assumption, low density separation (cluster) assumption.

**Ours Approach (SNTG) at a Glance**

- **Key idea:**
  - Explore more information in teacher than the target quality using the teacher graph.
  - Encourage features to form tighter clusters and keep the decision boundaries far away from data.
  - Doubly stochastic sampling algorithm to reduce computation cost.

**Comparisons of Graphs & Ablation Study**

- **Shortcomings of other graphs:**
  - (a) K-NN in input space $X$ (Weston et al. 2008): Low-level, cannot reflect semantic similarity.
  - (b) A pre-trained fixed graph in $Y$ : Cannot receive feedbacks from the model in the training.

**Experiments & Results**

- **SOTA semi-supervised classification:**
  1. MNIST: 10 labels (1.36 ± 0.75); 100 labels (0.66 ± 0.07)
  2. SVHN: 250 labels (4.29 ± 0.23); 500 labels (3.99 ± 0.24)
  3. CIFAR-10: 1000 labels (18.41 ± 0.52); 2000 labels (13.64 ± 0.32)

- **GAN-based generative SSL benefits from SNTG:**
  - Improved generation quality & diversity, without strange patterns.

- **Robustness to noisy labels:** SVHN with certain percentages of corrupted labels.

- **Visualization of embedding:** II model v.s. SNTG

Source code available at https://github.com/ximwei9322/SNTG