Graph Convolution for Semi-Supervised Classification: Improved Linear Separability and Out-of-Distribution Generalization

Aseem Baranwal, Kimon Fountoulakis*, Aukosh Jagannath*

Objective
- Study the effect of graph convolution on linear separability of a GMM
- Analyze generalization potential of the cross-entropy minimizer
- Conduct experiments in various settings to illustrate our results

Effect of Graph Convolution

Separability Results
- Without the graph, no hyperplane can separate a binary GMM if
  \[ \| \mu - \nu \|_1 = \mathcal{O}(\sigma) \]
- With graph convolution, this threshold changes to
  \[ \| \mu - \nu \| = \Theta \left( \frac{\sigma}{\sqrt{\|E\|}} \right) \]
- Without graph convolution, the loss is lower bounded
  \[ \text{Loss} \geq (2 \log 2) \Phi \left( -\frac{\| \mu - \nu \|}{2\sigma} \right) \]
- When the convolved data is separable, the loss is upper bounded
  \[ \text{Loss}(A, X) \leq C \exp \left( -d \| \mu - \nu \| \left( \frac{p - q}{p + q} \right) \right) \]

Generalization
- For any new dataset \( A, X \) with different \( n, p, q \), loss is bounded
  \[ \text{Loss}(A, X) \leq C \exp \left( -d \| \mu - \nu \| \left( \frac{p - q}{p + q} \right) \right) \]
- Loss increases with inter-class edge probability \( q \) (noisy graph)

When is Graph Convolution a Pitfall — Large \( q \)

- Large \( q \) implies higher loss — graph convolution hurts linear separability of data

Conclusions
- Graph convolution can transform linearly inseparable data into linearly separable data
- Graph convolution can be disadvantageous if the intra-class edge probability is close to the inter-class edge probability
- The learned classifier generalizes to out-of-distribution data

Future Work
- Extension to setting with multiple classes
- Properties of the optimal classifier for non-linearly separable data
- Understanding effects of graph convolution for highly non-linear models and deeper networks

*Supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). Cette recherche été financée par le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG), grant number [RGPIN-2019-04087, DGECR-2019-00147].

*Supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). Cette recherche été financée par le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG), grant number [RGPIN-2020-04597, DGECR-2020-00199].