Artificial Intelligence: Paper Critique

Maneet Singh (PhD15001)

Paper Title: Improved Deep Embedded Clustering with Local Structure Preservation

Authors: Xifeng Guo, Long Gao, Xinwang Liu, Jianping Yin

Venue: IJCAI 2017

Link: https://www.ijcai.org/proceedings/2017/0243.pdf

I. Summary

Most of the traditional clustering techniques such as K-means or spectral clustering perform clustering on the given data directly. For high dimensional data, this might result in undesirable results, since all dimensions are treated equally. In order to mitigate this problem, researchers have proposed techniques in order to transform the high dimensional data into a lower dimensional compact representation. This has given rise to the domain of *deep clustering*, wherein deep learning models are utilized to first obtain a lower dimensional representation of the data, which is then followed by some clustering technique. This paper presents a technique, *Improved Deep Embedded Clustering (IDEC)* for performing dimensionality reduction and clustering simultaneously. The proposed technique is built over Deep Embedded Clustering (DEC) [1], proposed earlier.

A. Deep Embedded Clustering

Deep Embedded Clustering (DEC) [1] uses an autoencoder and KL-divergence based loss for learning a low dimensional representation and performing clustering simultaneously. An autoencoder consists of en encoding and decoding layer - encoding layer is responsible for learning the representation, while the decoding layer reconstructs the input from the learned representation. The model thus minimizes the loss between the actual input and the reconstructed input, thereby facilitating *representative* feature learning. In DEC, the authors utilize a pre-trained autoencoder without its decoding layer. A KL-divergence based loss is added to the encoder, which is then trained for the task of dimensionality reduction and clustering by optimizing the following loss:

$$L_c = KL(P||Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}} \tag{1}$$

here, q_{ij} refers to the distance of sample i's representation from μ_j (mean of the j^{th} cluster). p_{ij} is the target distribution, defined in terms of q_{ij} as:

$$p_{ij} = \frac{q_{ij}^2 / \sum_i q_{ij}}{\sum_j (q_{ij}^2 / \sum_i q_{ij})} \tag{2}$$

The authors refer to this as a form of self-learning since the target distribution is a function of the sample's distances from different means. L_c is referred as the clustering loss, which is the key novelty presented in Deep Embedded Clustering. The main limitation of this technique was that it used only the encoder while fine-tuning the representation for clustering. This implies that the *representative* property of the autoencoder is not utilized for learning clustering-oriented features.

B. Improved Deep Embedded Clustering

In order to address the key limitation of Deep Embedded Clustering, the authors proposed Improved Deep Embedded Clustering. Instead of discarding the decoder (as in case of DEC), the authors optimize the reconstruction loss and the clustering loss together. The loss function of the Improved Deep Embedded Clustering technique can thus be written as:

$$L_{IDEC} = L_{AE} + L_C \tag{3}$$

The proposed technique learns *representative* features using an Autoencoder, and performs clustering on the learned representations simultaneously. This ensures that the model learns features which are representative of the data, while being useful for the task of clustering. At the time of optimization, the encoding layer receives gradient from the decoding layer and the clustering loss component.

Experimental analysis is performed on three datasets of MNIST, USPS, and Reuters-10K. Comparison has been performed with DEC and two other techniques. The model showcases improved performance in comparison to other techniques reported in the paper.

II. Limitations of the Proposed Technique

The key limitations of the proposed Improved Deep Embedded Clustering are as follows:

- A major limitation of the proposed technique is the pre-requisite knowledge of the number of clusters to be formed. Moreover, while an initial estimate is obtained using the number of clusters, the model is not even updated to increase/decrease the total clusters based on the feature representations.
- The proposed technique does not explicitly try to bring samples which are far from the cluster centers into one of the clusters. This might result in the model learning representations guided by the majority of samples only.
- There is not sufficient explanation or analysis of the clustering loss (L_C) that has been added to the autoencoder model. The authors have not expressed why this specific form of loss, taken from the DEC algorithm, would be optimal for such an architecture.

III. Suggestions for the Proposed Technique

Some suggestions to improve the proposed technique are as follows:

- An intuitive way of addressing the problem of fixed number of clusters is by incorporating a dynamic clustering based technique in the proposed model, where the number of clusters may also get updated. Any hierarchical clustering technique can also be applied to modify the number of clusters.
- In order to facilitate learning of representations which result in dense clusters, an additional distance-weighted term can be added in the proposed model which gives more weight to samples which are farther away from the centers.
- For addressing the third limitation, the authors should explicitly explain how the clustering loss helps in creation of better clusters. Specifically, how does the dynamic target distribution affect the clustering performance. Analysis should also be done with different clustering based loss in the autoencoder model as well.

References

[1] J. Xie, R. Girshick, and A. Farhadi. Unsupervised deep embedding for clustering analysis. In *Proceedings of International Conference on Machine Learning*, pages 478–487, 2016.