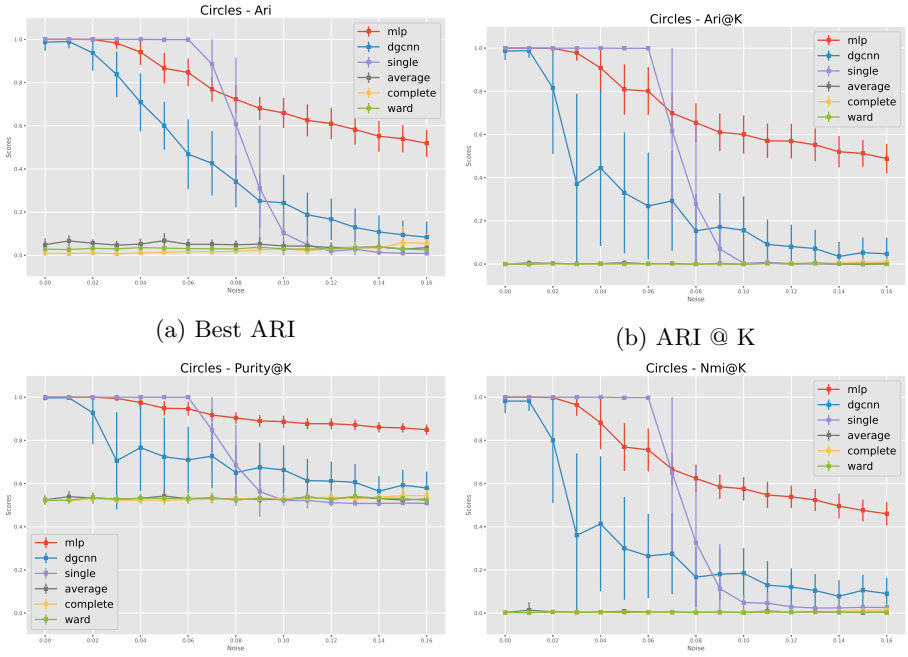
End-to-End Similarity Learning and Hierarchical clustering for unfixed size datasets - Supplementary Material

Leonardo Gigli, Santiago Velasco-Forero, Beatriz Marcotegui

February 2021

Appendix A. Supplementary Figures section Experiments



(a) Best ARI

(b) ARI @ K

(c) Purity @ K

(d) NMI @ K

Fig. 1: Robustness to noise of models on Circles. We compare trained models against classical methods as Single Linkage, Average Linkage, Complete Linkage and Ward's Method. The models used have been trained on a dataset without noise. Test sets used to measure scores contain 20 samples each. Plots show mean and standard deviation of scores obtained. During the experiments on this dataset MLP has shown a higher robustness to noise compared with DGCNN. Among classical methods only single linkage perform well on these samples.

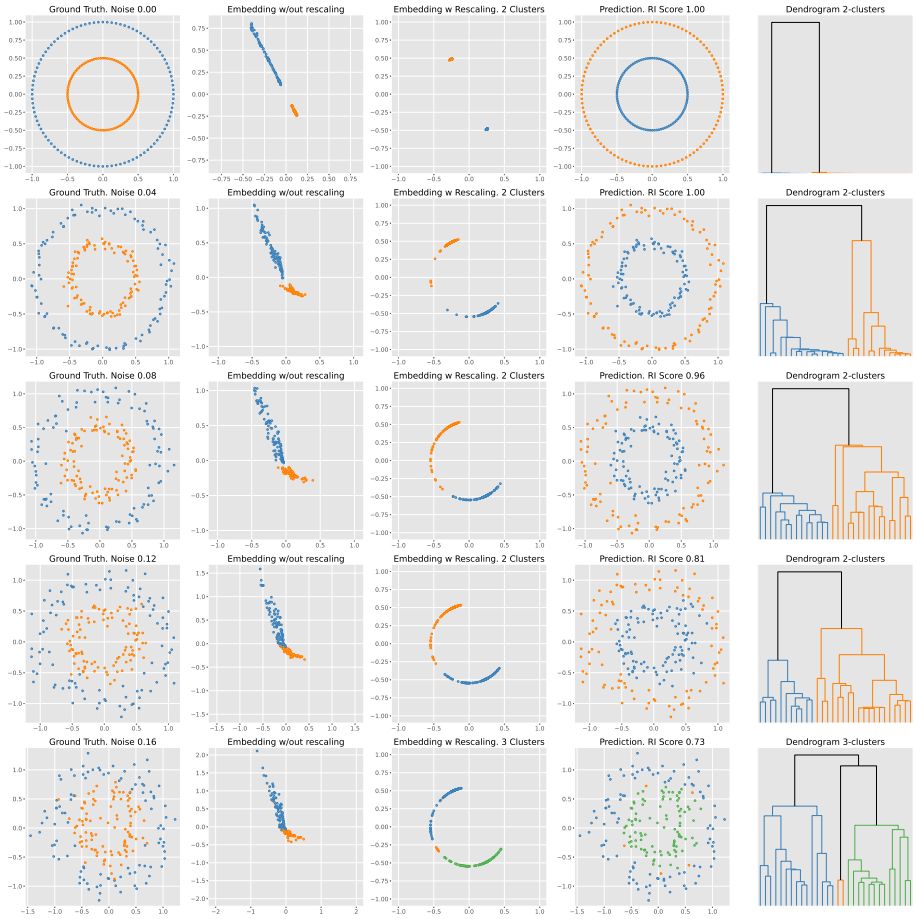


Fig. 2: Effect of noise on predictions in the circle database. The model used for prediction is a MLP trained using a dataset without noise. From top to bottom, each row is a case with an increasing level of noise. In the first column the input points, while in the second column we illustrate hidden features. Points are colored according to ground truth. The third column illustrates the hidden features after projection to Poincaré Disk. Fourth column shows predicted labels, while the fifth column shows associated dendrograms. Colors in the last three columns are assigned according to predicted labels.

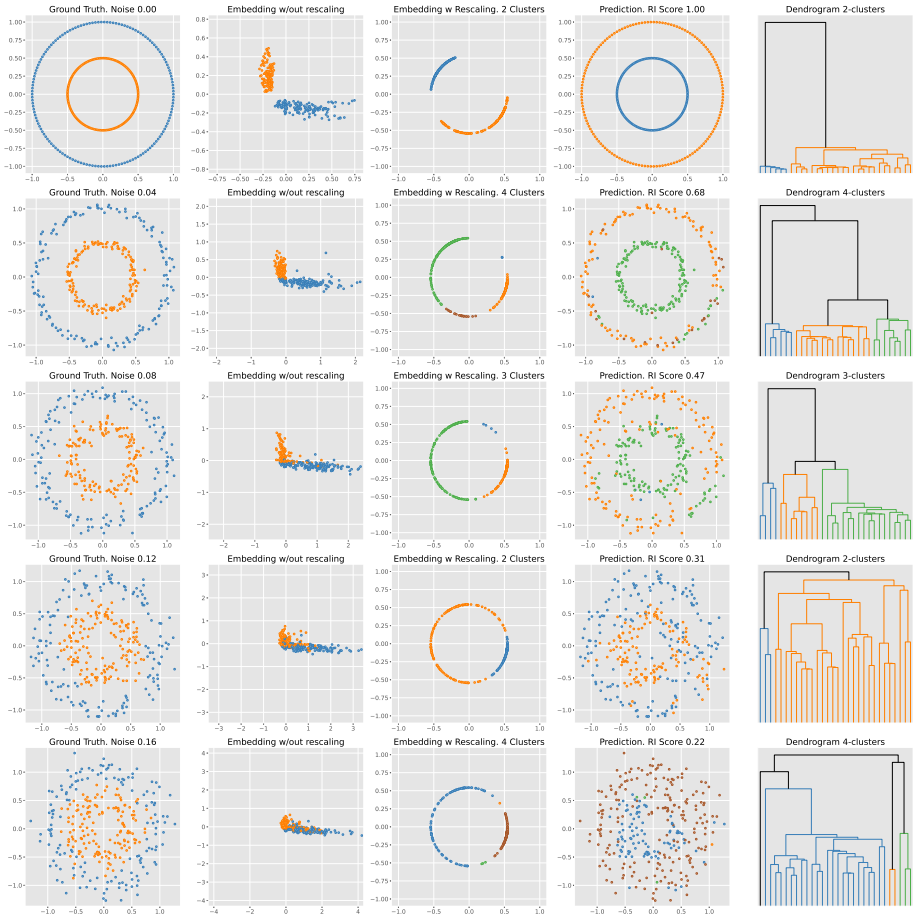
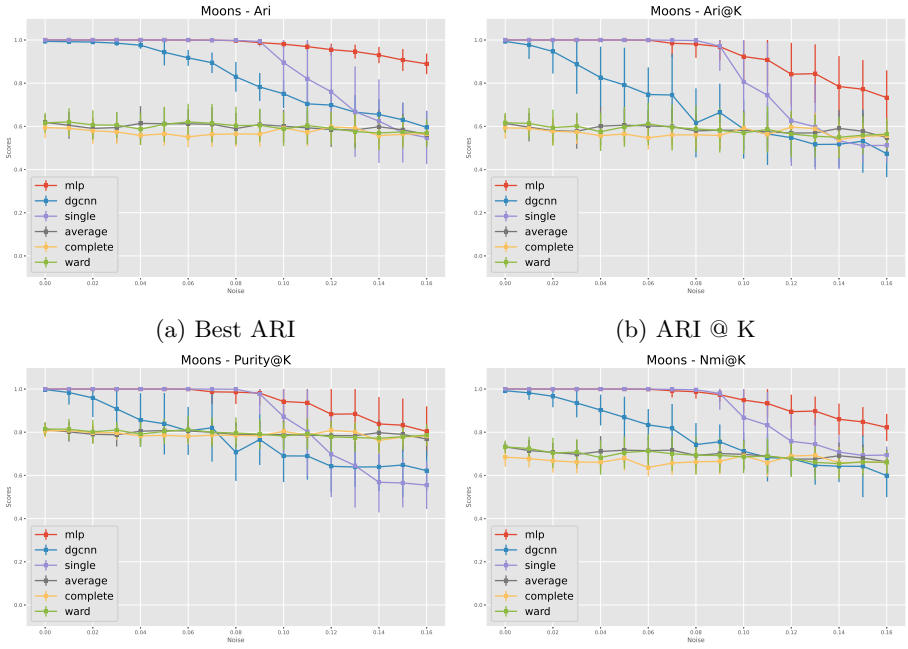


Fig. 3: Effect of noise on predictions in circle database. The model used for prediction is a DGCNN trained using a dataset without noise. From top to bottom, each row is a case with an increasing level of noise. In the first column the input points, while in the second column we illustrate hidden features. Points are colored according to ground truth. The third column illustrates the hidden features after projection to Poincaré Disk. Fourth column shows predicted labels, while the fifth column shows associated dendrograms. Colors in the last three columns are assigned according to predicted labels.



(a) Best ARI

(b) ARI @ K

(c) Purity @ K

(d) NMI @ K

Fig. 4: Robustness to noise of models on Moons. We compare trained models against classical methods as Single Linkage, Average Linkage, Complete Linkage, Ward's Method. The models used have been trained on a dataset without noise. Test sets used to measure scores contain 20 samples each. Plots show mean and standard deviation of scores obtained. During the experiments on these datasets MLP has shown a higher robustness to noise compared with the other models.

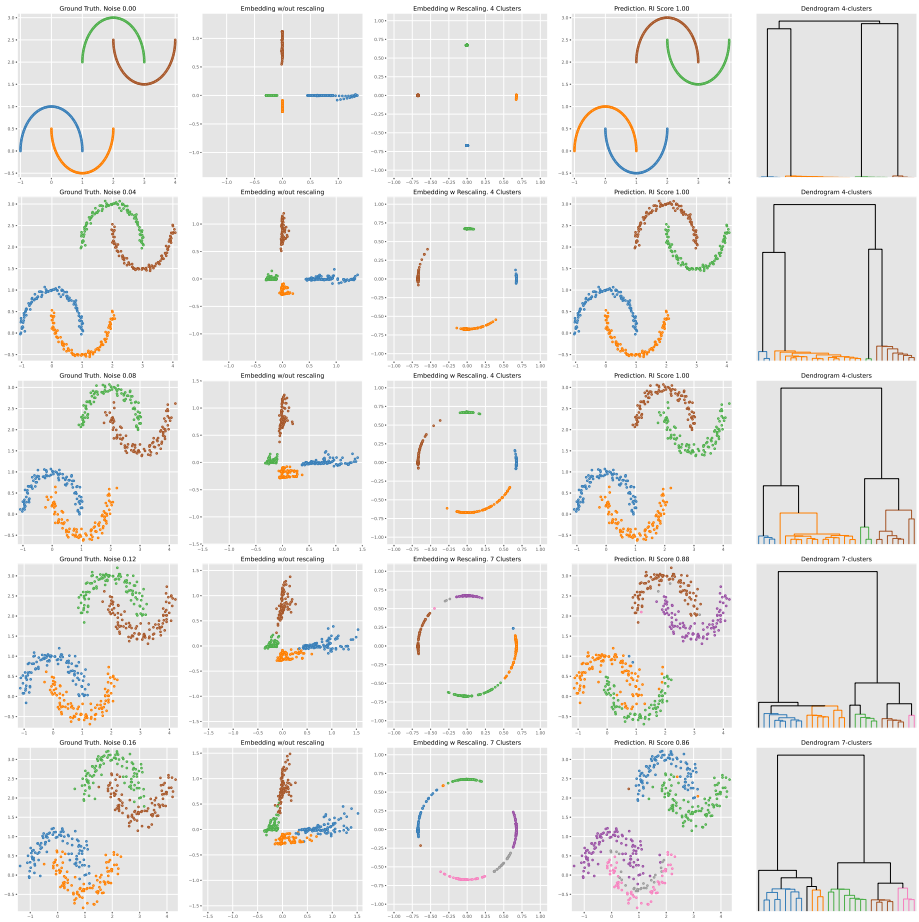


Fig. 5: Moons dataset. The model used is a MLP trained on samples without noise. From top to bottom, each row is a case with an increasing level of noise. In the first column the input points, while in the second column we illustrate hidden features. Points are colored according to ground truth. The third column illustrates the hidden features after projection to Poincaré Disk. Fourth column shows predicted labels, while the fifth column shows associated dendrograms. Colors in the last three columns are assigned according to predicted labels.

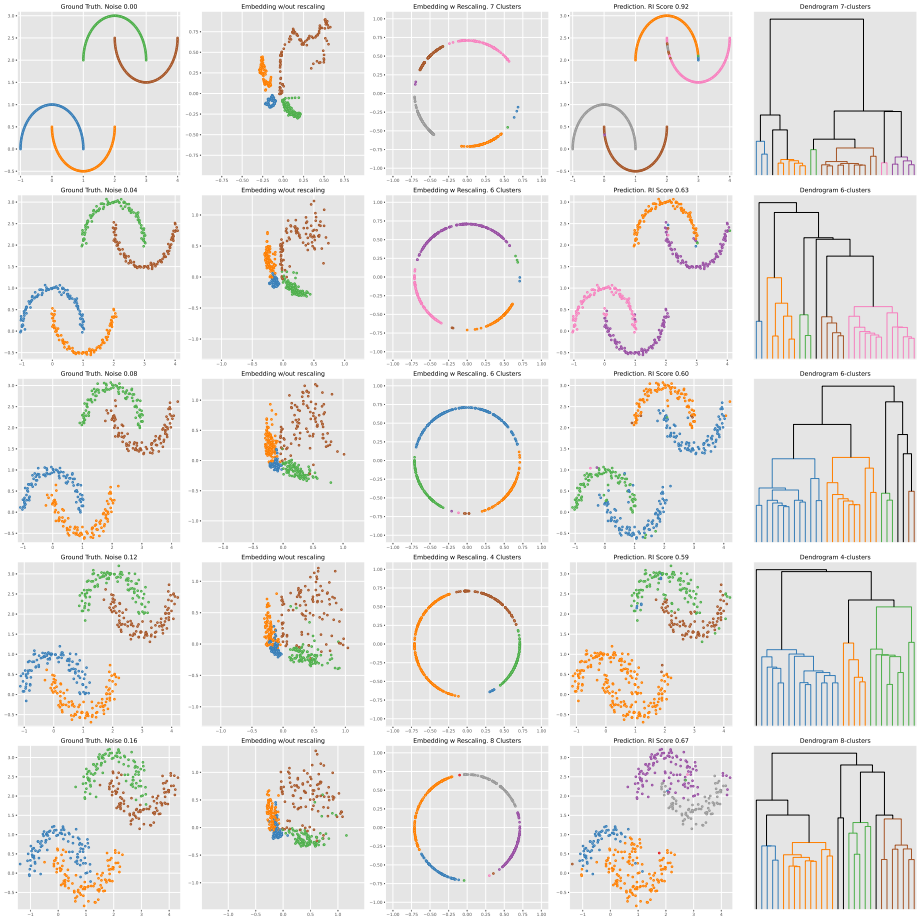
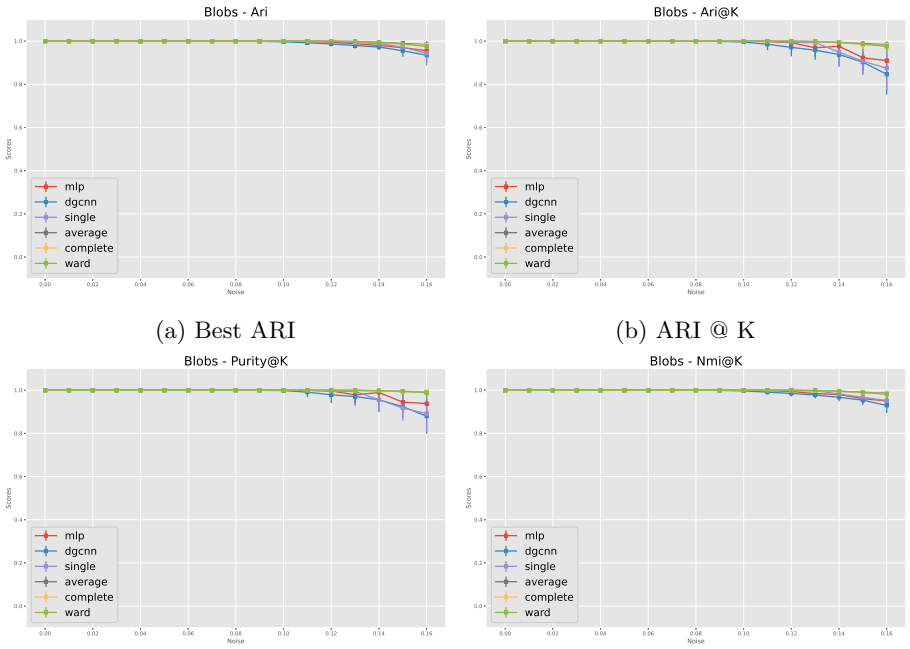


Fig. 6: Moons dataset. The model used for prediction is a DGCNN trained on samples without noise. From top to bottom, each row is a case with an increasing level of noise. In the first column the input points, while in the second column we illustrate hidden features. Points are colored according to ground truth. The third column illustrates the hidden features after projection to Poincaré Disk. Fourth column shows predicted labels, while the fifth column shows associated dendrograms. Colors in the last three columns are assigned according to predicted labels.



(a) Best ARI

(b) ARI @ K

(c) Purity @ K

(d) NMI @ K

Fig. 7: Robustness to noise of models on Blobs. We compare trained models against classical methods as Single Linkage, Average Linkage, Complete Linkage and Ward's Method. The models used have been trained on a dataset with Gaussian's standard deviation fixed at 0.08. Test sets used to measure scores contain 20 samples each. Plots show mean and standard deviation of scores obtained. During the experiments on these datasets MLP has shown a higher robustness to noise compared with DGCNN. In this case classical methods show better performances compared to trained models.

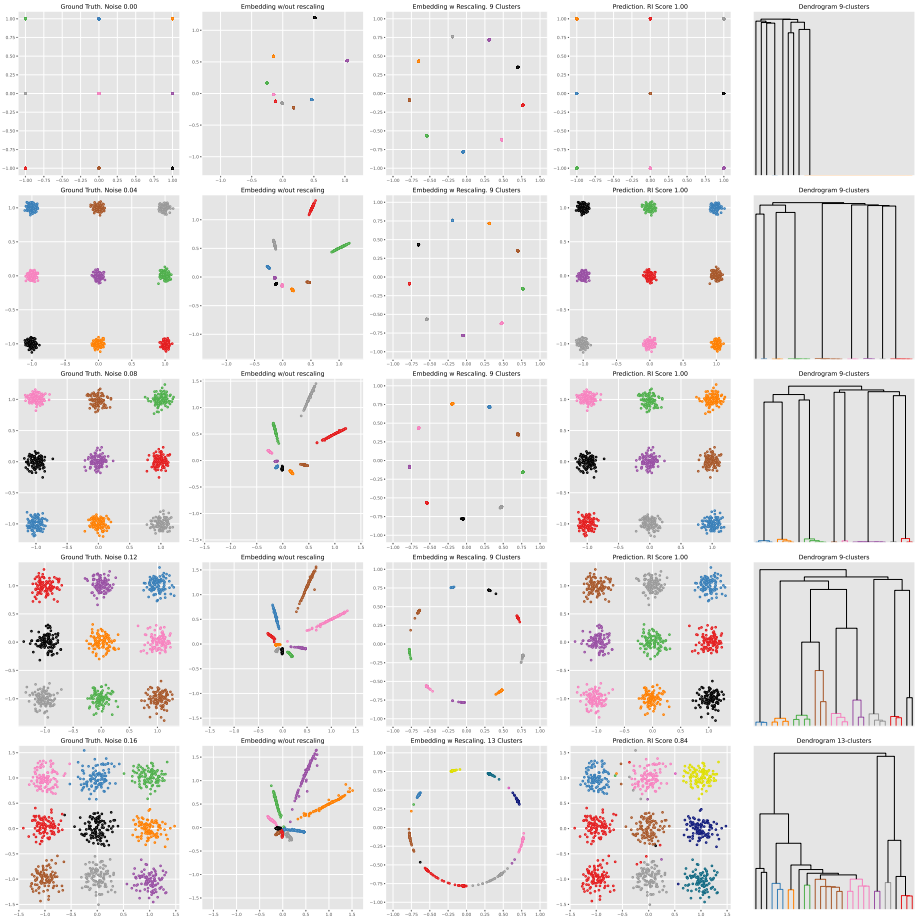


Fig. 8: Blobs dataset. The model used for prediction is a MLP trained on samples with standard deviation value at 0.08. From top to bottom, each row is a case with an increasing value of standard deviation. In the first column the input points, while in the second column we illustrate hidden features. Points are colored according to ground truth. The third column illustrates the hidden features after projection to Poincaré Disk. Forth column shows predicted labels, while the fifth column show associated dendrograms. Colors in the last three columns are assigned according to predicted labels.

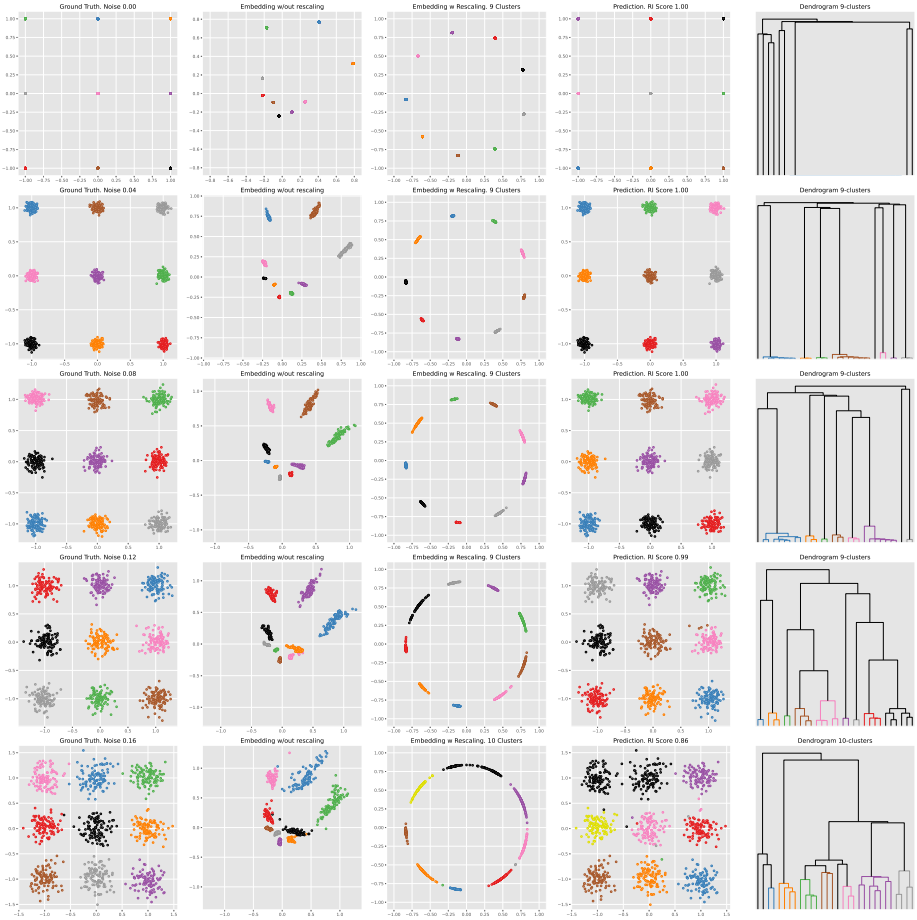
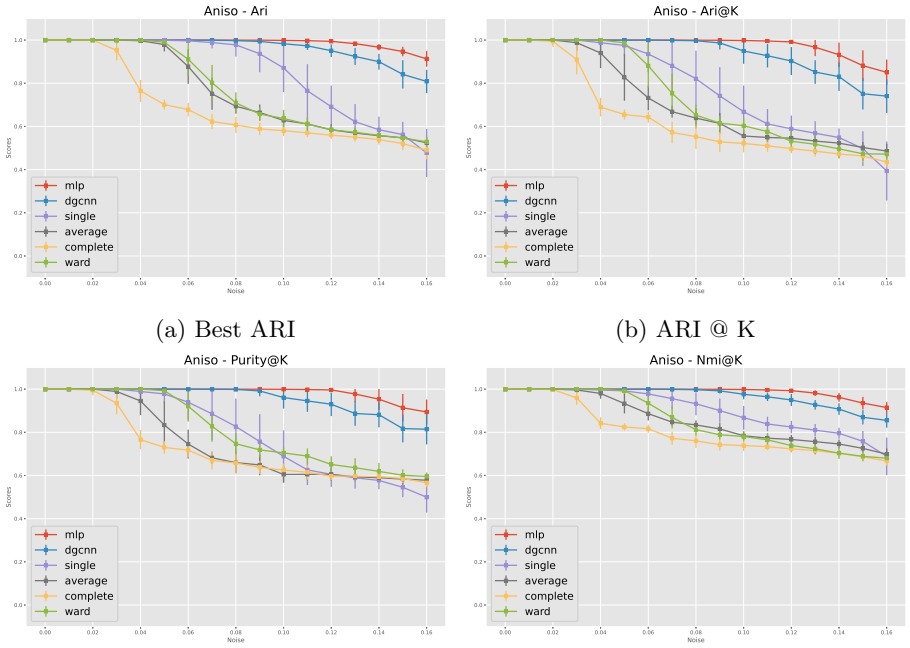


Fig. 9: Blobs dataset. The model used for prediction is a DGCNN trained on samples with standard deviation value at 0.08. From top to bottom, each row is a case with an increasing value of standard deviation. In the first column the input points, while in the second column we illustrate hidden features. Points are coloured according to the ground truth labels. The third column illustrates the hidden features after projection to Poincaré Disk. Fourth column shows predicted labels, while the fifth column show associated dendrograms. Colors in the last three columns are assigned according to predicted labels.



(a) Best ARI

(b) ARI @ K

(c) Purity @ K

(d) NMI @ K

Fig. 10: Robustness to noise of models on Anisotropic dataset. We compare trained models against classical methods as Single Linkage, Average Linkage, Complete Linkage and Ward's Method. The models used have been trained on a dataset with Gaussian's standard deviation fixed at 0.08. Test sets used to measure scores contain 20 samples each. Plots show mean and standard deviation of scores obtained. During the experiments on these datasets MLP has shown a higher robustness to noise compared with DGCNN.

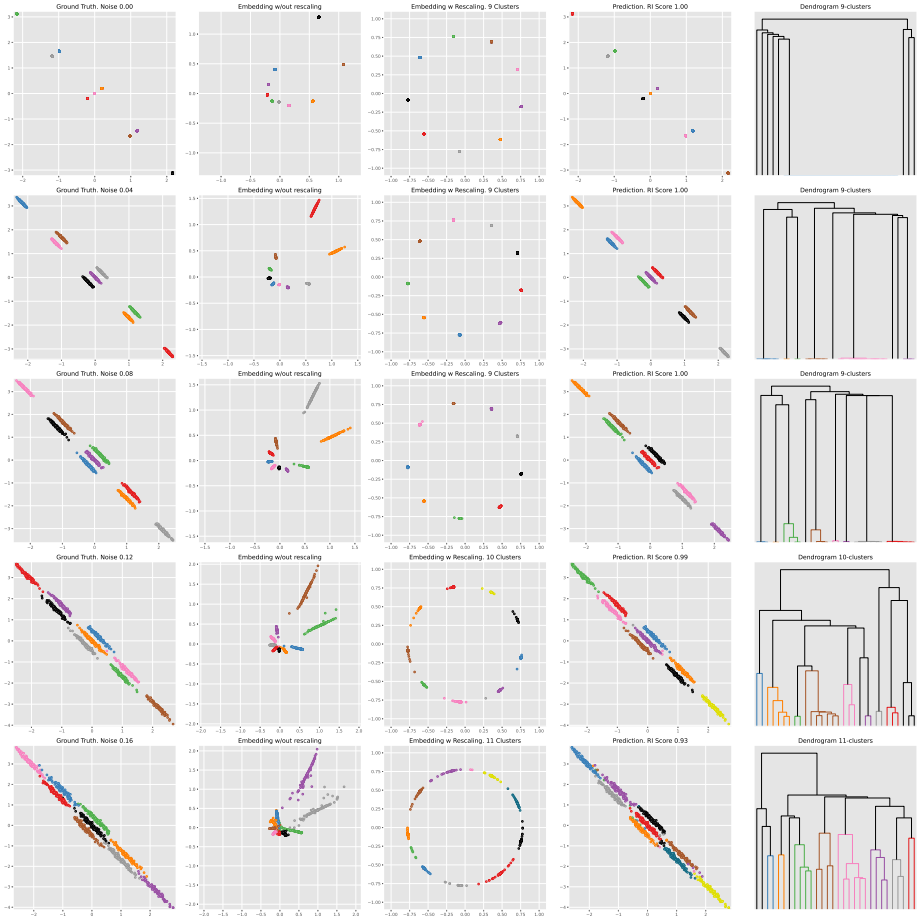


Fig. 11: Anisotropic dataset. The model used for prediction is a MLP trained on samples with standard deviation value at 0.08. From top to bottom, each rows is a case with an increasing value of standard deviation. In the first column the input points, while in the second column we illustrate hidden features. Points are colored based on ground truth. The third column illustrate the hidden features after projection to Poincaré Disk. Forth column shows predicted labels, while the fifth column show associated dendrograms. Colors in the last three columns are assigned according to predicted labels.

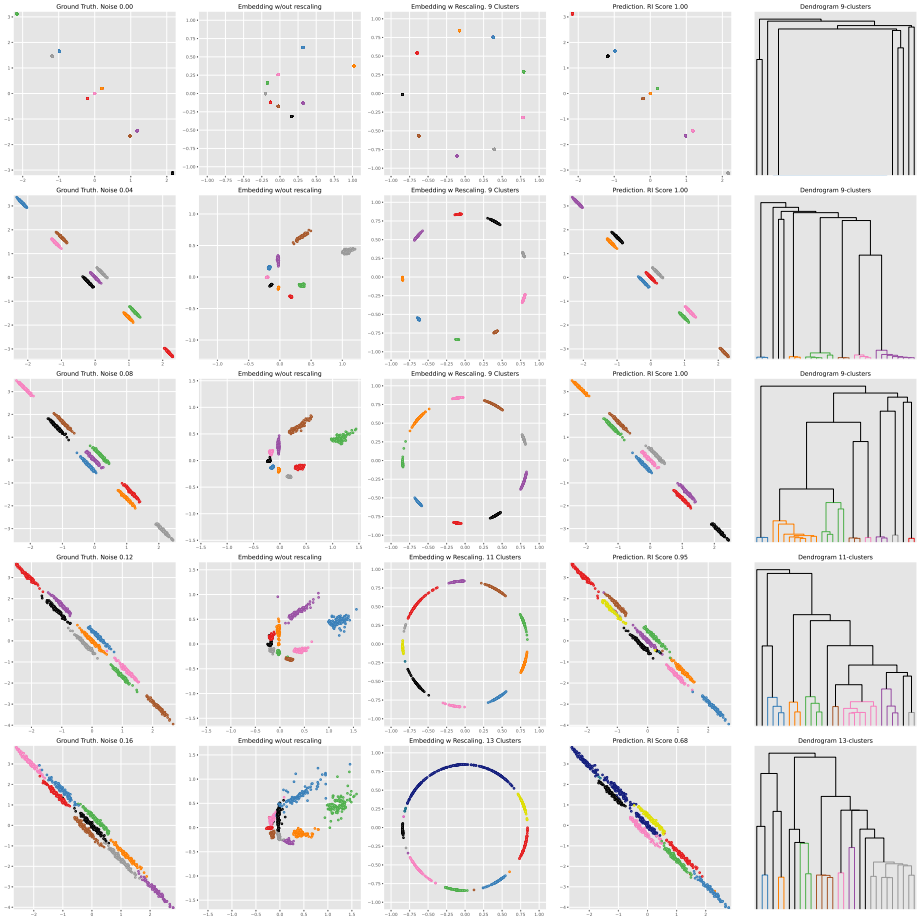


Fig. 12: Anisotropic dataset. The model used for prediction is a DGCNN trained on samples with standard deviation value at 0.08. From top to bottom, each rows is a case with an increasing value of standard deviation. In the first column the input points, while in the second column we illustrate hidden features. Points are colored according to ground truth. The third column illustrate the hidden features after projection to Poincaré Disk. Forth column shows predicted labels, while the fifth column show associated dendrograms. Colors in the last three columns are assigned according to predicted labels.