Tommy Inouye
Uni: ti2181
12/8/2014
Assignment 4

Conclusion

To analyze the effect of max deviation on the Macro F1 score, I created a simple for loop

that iterated through max deviation values from 0-10 at .1 intervals. The F1 Score peaked when

the max deviation was between .9 and 1.6. The best results are shown below and the results for

the whole iteration are shown at the end of this file.

The results shown below reflect that instance. The observed F1 score was around .79524.

That was 0.07339 more than the unpruned dataset and was a 10% increase. The graph also

shows how the learner after pruning not only performs better but is able to finish by evaluating

much less nodes. The peak of the graph is found in significantly less nodes in the pruned graph

compared to the unpruned graph, and is overall performs much better in all accounts.

So in conclusion, I can with very high confidence state that I was successful in

implementing the pruning function for my learner. There is no instance in my 100 iterations

where any pruned iterations performed worse than the unpruned one. In addition, with the

right max deviation, the pruned decision tree increased in accuracy by 10% and generated

much less nodes.
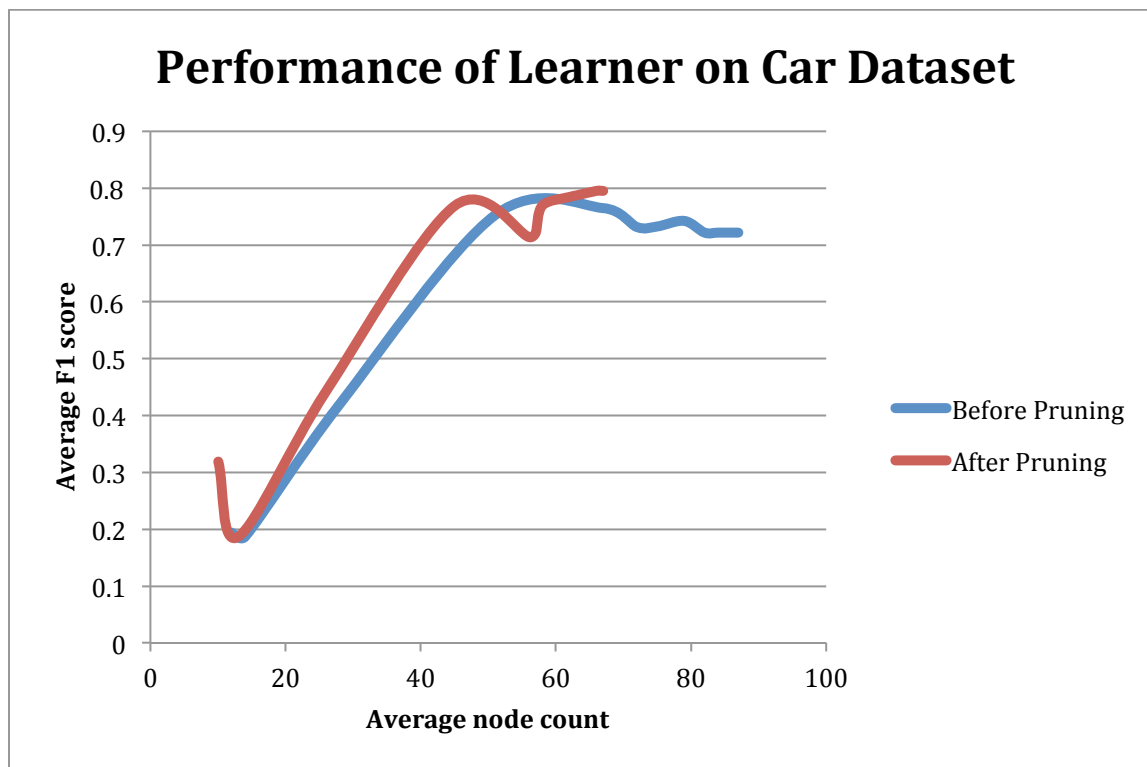
Analysis

RESULTS FOR Max Deviation=1.6
Macro F1 score on test data: 0.79524
Train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
Average node count on test dataset: ['10.00', '13.00', '26.00', '45.00', '56.00', '57.00', '58.00', '62.00', '63.00', '66.00', '67.00']
Average F1 score on test dataset: ['0.31952', '0.18690', '0.44278', '0.76876', '0.71406', '0.72897', '0.77700', '0.78450', '0.78715', '0.79524', '0.79524']

Graph of data

####RESULTS FOR ALL MAXDEVIATIONS 0-10########
Before Pruning
========================
Macro F1 score on test data: 0.72185
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['12.00', '14.00', '28.00', '51.00', '67.00', '72.00', '75.00', '79.00', '82.00', '84.00', '87.00']
average F1 score on test dataset: ['0.19427', '0.18690', '0.41988', '0.75262', '0.76495', '0.73162', '0.73269', '0.74247', '0.72185', '0.72185', '0.72185']

After Pruning
========================
testing maxdeviation=0.0
Macro F1 score on test data: 0.72185
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['12.00', '14.00', '28.00', '51.00', '67.00', '72.00', '75.00', '79.00', '82.00', '84.00', '87.00']
average F1 score on test dataset: ['0.19427', '0.18690', '0.41988', '0.75262', '0.76495', '0.73162', '0.73269', '0.74247', '0.72185', '0.72185', '0.72185']
========================
testing maxdeviation=0.1
Macro F1 score on test data: 0.77030
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['11.00', '13.00', '26.00', '45.00', '56.00', '58.00', '60.00', '64.00', '65.00', '68.00', '69.00']
average F1 score on test dataset: ['0.18852', '0.18690', '0.44278', '0.76876', '0.71406', '0.73187', '0.77700', '0.78450', '0.76211', '0.77030', '0.77030']
========================
testing maxdeviation=0.2
Macro F1 score on test data: 0.77030
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['11.00', '13.00', '26.00', '45.00', '56.00', '58.00', '60.00', '64.00', '65.00', '68.00', '69.00']
average F1 score on test dataset: ['0.18852', '0.18690', '0.44278', '0.76876', '0.71406', '0.73187', '0.77700', '0.78450', '0.76211', '0.77030', '0.77030']
========================
testing maxdeviation=0.3
Macro F1 score on test data: 0.77030
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['11.00', '13.00', '26.00', '45.00', '56.00', '58.00', '60.00', '64.00', '65.00', '68.00', '69.00']
average F1 score on test dataset: ['0.18852', '0.18690', '0.44278', '0.76876', '0.71406', '0.73187', '0.77700', '0.78450', '0.76211', '0.77030', '0.77030']
========================
testing maxdeviation=0.4
Macro F1 score on test data: 0.77030
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['11.00', '13.00', '26.00', '45.00', '56.00', '58.00', '60.00', '64.00', '65.00', '68.00', '69.00']
average F1 score on test dataset: ['0.18852', '0.18690', '0.44278', '0.76876', '0.71406', '0.73187', '0.77700', '0.78450', '0.76211', '0.77030', '0.77030']
========================
testing maxdeviation=0.5
Macro F1 score on test data: 0.77030

train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['11.00', '13.00', '26.00', '45.00', '56.00', '58.00', '60.00', '64.00', '65.00', '68.00', '69.00']
average F1 score on test dataset: ['0.18852', '0.18690', '0.44278', '0.76876', '0.71406', '0.73187', '0.77700', '0.78450', '0.76211', '0.77030', '0.77030']
========================
testing maxdeviation=0.6
Macro F1 score on test data: 0.77030
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['11.00', '13.00', '26.00', '45.00', '56.00', '58.00', '60.00', '64.00', '65.00', '68.00', '69.00']
average F1 score on test dataset: ['0.18852', '0.18690', '0.44278', '0.76876', '0.71406', '0.73187', '0.77700', '0.78450', '0.76211', '0.77030', '0.77030']
========================
testing maxdeviation=0.7
Macro F1 score on test data: 0.77030
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['11.00', '13.00', '26.00', '45.00', '56.00', '58.00', '60.00', '64.00', '65.00', '68.00', '69.00']
average F1 score on test dataset: ['0.18852', '0.18690', '0.44278', '0.76876', '0.71406', '0.73187', '0.77700', '0.78450', '0.76211', '0.77030', '0.77030']
========================
testing maxdeviation=0.8
Macro F1 score on test data: 0.77030
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['11.00', '13.00', '26.00', '45.00', '56.00', '57.00', '59.00', '63.00', '65.00', '68.00', '69.00']
average F1 score on test dataset: ['0.18852', '0.18690', '0.44278', '0.76876', '0.71406', '0.72897', '0.77376', '0.78113', '0.76211', '0.77030', '0.77030']
========================
testing maxdeviation=0.9
Macro F1 score on test data: 0.79524
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['11.00', '13.00', '26.00', '45.00', '56.00', '57.00', '58.00', '62.00', '63.00', '66.00', '67.00']
average F1 score on test dataset: ['0.18852', '0.18690', '0.44278', '0.76876', '0.71406', '0.72897', '0.77700', '0.78450', '0.78715', '0.79524', '0.79524']
========================
testing maxdeviation=1.0
Macro F1 score on test data: 0.79524
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['11.00', '13.00', '26.00', '45.00', '56.00', '57.00', '58.00', '62.00', '63.00', '66.00', '67.00']
average F1 score on test dataset: ['0.18852', '0.18690', '0.44278', '0.76876', '0.71406',

'0.72897', '0.77700', '0.78450', '0.78715', '0.79524', '0.79524']
========================
testing maxdeviation=1.1
Macro F1 score on test data: 0.79524
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['11.00', '13.00', '26.00', '45.00', '56.00', '57.00', '58.00', '62.00', '63.00', '66.00', '67.00']
average F1 score on test dataset: ['0.18852', '0.18690', '0.44278', '0.76876', '0.71406', '0.72897', '0.77700', '0.78450', '0.78715', '0.79524', '0.79524']
========================
testing maxdeviation=1.2
Macro F1 score on test data: 0.79524
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['11.00', '13.00', '26.00', '45.00', '56.00', '57.00', '58.00', '62.00', '63.00', '66.00', '67.00']
average F1 score on test dataset: ['0.18852', '0.18690', '0.44278', '0.76876', '0.71406', '0.72897', '0.77700', '0.78450', '0.78715', '0.79524', '0.79524']
========================
testing maxdeviation=1.3
Macro F1 score on test data: 0.79524
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['11.00', '13.00', '26.00', '45.00', '56.00', '57.00', '58.00', '62.00', '63.00', '66.00', '67.00']
average F1 score on test dataset: ['0.18852', '0.18690', '0.44278', '0.76876', '0.71406', '0.72897', '0.77700', '0.78450', '0.78715', '0.79524', '0.79524']
========================
testing maxdeviation=1.4
Macro F1 score on test data: 0.79524
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['10.00', '13.00', '26.00', '45.00', '56.00', '57.00', '58.00', '62.00', '63.00', '66.00', '67.00']
average F1 score on test dataset: ['0.31952', '0.18690', '0.44278', '0.76876', '0.71406', '0.72897', '0.77700', '0.78450', '0.78715', '0.79524', '0.79524']
========================
testing maxdeviation=1.5
Macro F1 score on test data: 0.79524
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['10.00', '13.00', '26.00', '45.00', '56.00', '57.00', '58.00', '62.00', '63.00', '66.00', '67.00']
average F1 score on test dataset: ['0.31952', '0.18690', '0.44278', '0.76876', '0.71406', '0.72897', '0.77700', '0.78450', '0.78715', '0.79524', '0.79524']
========================
testing maxdeviation=1.6
Macro F1 score on test data: 0.79524

train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['10.00', '13.00', '26.00', '45.00', '56.00', '57.00', '58.00', '62.00', '63.00', '66.00', '67.00']
average F1 score on test dataset: ['0.31952', '0.18690', '0.44278', '0.76876', '0.71406', '0.72897', '0.77700', '0.78450', '0.78715', '0.79524', '0.79524']
========================
testing maxdeviation=1.7
Macro F1 score on test data: 0.79524
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['10.00', '13.00', '26.00', '45.00', '56.00', '57.00', '58.00', '62.00', '63.00', '66.00', '67.00']
average F1 score on test dataset: ['0.31952', '0.18690', '0.44278', '0.76876', '0.71406', '0.72897', '0.77700', '0.78450', '0.78715', '0.79524', '0.79524']
========================
testing maxdeviation=1.8
Macro F1 score on test data: 0.79524
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['10.00', '13.00', '26.00', '45.00', '56.00', '57.00', '58.00', '62.00', '63.00', '66.00', '67.00']
average F1 score on test dataset: ['0.31952', '0.18690', '0.44278', '0.76876', '0.71406', '0.72897', '0.77700', '0.78450', '0.78715', '0.79524', '0.79524']
========================
testing maxdeviation=1.9
Macro F1 score on test data: 0.79524
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['10.00', '12.00', '26.00', '45.00', '56.00', '57.00', '58.00', '62.00', '63.00', '66.00', '67.00']
average F1 score on test dataset: ['0.31952', '0.28636', '0.44278', '0.76876', '0.71406', '0.72897', '0.77700', '0.78450', '0.78715', '0.79524', '0.79524']
========================
testing maxdeviation=2.0
Macro F1 score on test data: 0.79524
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['10.00', '12.00', '26.00', '45.00', '56.00', '57.00', '58.00', '62.00', '63.00', '66.00', '67.00']
average F1 score on test dataset: ['0.31952', '0.28636', '0.44278', '0.76876', '0.71406', '0.72897', '0.77700', '0.78450', '0.78715', '0.79524', '0.79524']
========================
testing maxdeviation=2.1
Macro F1 score on test data: 0.78413
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '11.00', '21.00', '41.00', '49.00', '53.00', '55.00', '58.00', '59.00', '62.00', '64.00']
average F1 score on test dataset: ['0.32310', '0.30697', '0.49506', '0.74194', '0.70121',

'0.71424', '0.76473', '0.76885', '0.77558', '0.78401', '0.78413']
=======================
testing maxdeviation=2.2
Macro F1 score on test data: 0.78413
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '11.00', '21.00', '41.00', '49.00', '53.00', '55.00', '58.00', '59.00', '62.00', '64.00']
average F1 score on test dataset: ['0.32310', '0.30697', '0.49506', '0.74194', '0.70121', '0.71424', '0.76473', '0.76885', '0.77558', '0.78401', '0.78413']
=======================
testing maxdeviation=2.3
Macro F1 score on test data: 0.78167
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '11.00', '21.00', '41.00', '49.00', '53.00', '55.00', '57.00', '58.00', '61.00', '63.00']
average F1 score on test dataset: ['0.32310', '0.30697', '0.49506', '0.74194', '0.70121', '0.71424', '0.76473', '0.76635', '0.77306', '0.78150', '0.78167']
=======================
testing maxdeviation=2.4
Macro F1 score on test data: 0.78167
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '11.00', '21.00', '41.00', '49.00', '53.00', '55.00', '57.00', '58.00', '61.00', '63.00']
average F1 score on test dataset: ['0.32310', '0.30697', '0.49506', '0.74194', '0.70121', '0.71424', '0.76473', '0.76635', '0.77306', '0.78150', '0.78167']
=======================
testing maxdeviation=2.5
Macro F1 score on test data: 0.78167
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '11.00', '21.00', '41.00', '49.00', '53.00', '55.00', '57.00', '58.00', '61.00', '63.00']
average F1 score on test dataset: ['0.32310', '0.30697', '0.49506', '0.74194', '0.70121', '0.71424', '0.76473', '0.76635', '0.77306', '0.78150', '0.78167']
=======================
testing maxdeviation=2.6
Macro F1 score on test data: 0.78167
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '11.00', '21.00', '41.00', '49.00', '53.00', '55.00', '57.00', '58.00', '61.00', '63.00']
average F1 score on test dataset: ['0.32310', '0.30697', '0.49506', '0.74194', '0.70121', '0.71424', '0.76473', '0.76635', '0.77306', '0.78150', '0.78167']
=======================
testing maxdeviation=2.7
Macro F1 score on test data: 0.78167

train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '11.00', '21.00', '40.00', '48.00', '53.00', '55.00', '57.00', '58.00', '61.00', '63.00']
average F1 score on test dataset: ['0.32310', '0.30697', '0.49506', '0.73932', '0.69865', '0.71424', '0.76473', '0.76635', '0.77306', '0.78150', '0.78167']
========================
testing maxdeviation=2.8
Macro F1 score on test data: 0.78167
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '11.00', '21.00', '40.00', '48.00', '53.00', '55.00', '57.00', '58.00', '61.00', '63.00']
average F1 score on test dataset: ['0.32310', '0.30697', '0.49506', '0.73932', '0.69865', '0.71424', '0.76473', '0.76635', '0.77306', '0.78150', '0.78167']
========================
testing maxdeviation=2.9
Macro F1 score on test data: 0.78167
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '11.00', '21.00', '40.00', '48.00', '53.00', '55.00', '57.00', '58.00', '61.00', '63.00']
average F1 score on test dataset: ['0.32310', '0.30697', '0.49506', '0.73932', '0.69865', '0.71424', '0.76473', '0.76635', '0.77306', '0.78150', '0.78167']
========================
testing maxdeviation=3.0
Macro F1 score on test data: 0.78167
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '11.00', '21.00', '40.00', '47.00', '52.00', '54.00', '57.00', '58.00', '61.00', '63.00']
average F1 score on test dataset: ['0.32310', '0.30697', '0.49506', '0.73932', '0.69865', '0.71424', '0.76473', '0.76635', '0.77306', '0.78150', '0.78167']
========================
testing maxdeviation=3.1
Macro F1 score on test data: 0.78866
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '44.00', '49.00', '50.00', '53.00', '53.00', '56.00', '58.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69751', '0.71303', '0.76384', '0.76549', '0.77978', '0.78820', '0.78866']
========================
testing maxdeviation=3.2
Macro F1 score on test data: 0.78866
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '44.00', '49.00', '50.00', '53.00', '53.00', '56.00', '58.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69751',

'0.71303', '0.76384', '0.76549', '0.77978', '0.78820', '0.78866']
========================
testing maxdeviation=3.3
Macro F1 score on test data: 0.78866
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '44.00', '49.00', '50.00', '53.00', '53.00', '56.00', '58.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69751', '0.71303', '0.76384', '0.76549', '0.77978', '0.78820', '0.78866']
========================
testing maxdeviation=3.4
Macro F1 score on test data: 0.78866
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '44.00', '49.00', '50.00', '53.00', '53.00', '56.00', '58.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69751', '0.71303', '0.76384', '0.76549', '0.77978', '0.78820', '0.78866']
========================
testing maxdeviation=3.5
Macro F1 score on test data: 0.78866
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '44.00', '49.00', '50.00', '53.00', '53.00', '56.00', '58.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69751', '0.71303', '0.76384', '0.76549', '0.77978', '0.78820', '0.78866']
========================
testing maxdeviation=3.6
Macro F1 score on test data: 0.78866
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '44.00', '49.00', '50.00', '53.00', '53.00', '56.00', '58.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69751', '0.71303', '0.76384', '0.76549', '0.77978', '0.78820', '0.78866']
========================
testing maxdeviation=3.7
Macro F1 score on test data: 0.78866
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '44.00', '49.00', '50.00', '53.00', '53.00', '56.00', '58.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69751', '0.71303', '0.76384', '0.76549', '0.77978', '0.78820', '0.78866']
========================
testing maxdeviation=3.8
Macro F1 score on test data: 0.78866

train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '44.00', '49.00', '50.00', '53.00', '53.00', '56.00', '58.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69751', '0.71303', '0.76384', '0.76549', '0.77978', '0.78820', '0.78866']
=======================
testing maxdeviation=3.9
Macro F1 score on test data: 0.78866
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '44.00', '49.00', '50.00', '53.00', '53.00', '56.00', '58.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69751', '0.71303', '0.76384', '0.76549', '0.77978', '0.78820', '0.78866']
=======================
testing maxdeviation=4.0
Macro F1 score on test data: 0.78866
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '44.00', '49.00', '50.00', '53.00', '53.00', '56.00', '58.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69751', '0.71303', '0.76384', '0.76549', '0.77978', '0.78820', '0.78866']
=======================
testing maxdeviation=4.1
Macro F1 score on test data: 0.78167
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '44.00', '48.00', '48.00', '50.00', '50.00', '53.00', '55.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69751', '0.70477', '0.75860', '0.75860', '0.77306', '0.78150', '0.78167']
=======================
testing maxdeviation=4.2
Macro F1 score on test data: 0.78167
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '44.00', '48.00', '48.00', '50.00', '50.00', '53.00', '55.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69751', '0.70477', '0.75860', '0.75860', '0.77306', '0.78150', '0.78167']
=======================
testing maxdeviation=4.3
Macro F1 score on test data: 0.78167
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '44.00', '48.00', '48.00', '50.00', '50.00', '53.00', '55.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69751',

'0.70477', '0.75860', '0.75860', '0.77306', '0.78150', '0.78167']
========================
testing maxdeviation=4.4
Macro F1 score on test data: 0.78167
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '44.00', '48.00', '48.00', '50.00', '50.00', '53.00', '55.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69751', '0.70477', '0.75860', '0.75860', '0.77306', '0.78150', '0.78167']
========================
testing maxdeviation=4.5
Macro F1 score on test data: 0.78167
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '44.00', '48.00', '48.00', '50.00', '50.00', '53.00', '55.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69751', '0.70477', '0.75860', '0.75860', '0.77306', '0.78150', '0.78167']
========================
testing maxdeviation=4.6
Macro F1 score on test data: 0.78167
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '44.00', '48.00', '48.00', '50.00', '50.00', '53.00', '55.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69751', '0.70477', '0.75860', '0.75860', '0.77306', '0.78150', '0.78167']
========================
testing maxdeviation=4.7
Macro F1 score on test data: 0.78167
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '44.00', '48.00', '48.00', '50.00', '50.00', '53.00', '55.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69751', '0.70477', '0.75860', '0.75860', '0.77306', '0.78150', '0.78167']
========================
testing maxdeviation=4.8
Macro F1 score on test data: 0.77067
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '43.00', '47.00', '47.00', '49.00', '49.00', '52.00', '54.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69468', '0.70196', '0.74762', '0.74762', '0.76205', '0.77086', '0.77067']
========================
testing maxdeviation=4.9
Macro F1 score on test data: 0.77067

train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '43.00', '47.00', '47.00', '49.00', '49.00', '52.00', '54.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69468', '0.70196', '0.74762', '0.74762', '0.76205', '0.77086', '0.77067']
=========================
testing maxdeviation=5.0
Macro F1 score on test data: 0.77067
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '43.00', '47.00', '47.00', '49.00', '49.00', '52.00', '54.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69468', '0.70196', '0.74762', '0.74762', '0.76205', '0.77086', '0.77067']
=========================
testing maxdeviation=5.1
Macro F1 score on test data: 0.77067
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '43.00', '45.00', '45.00', '47.00', '47.00', '50.00', '52.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69468', '0.70196', '0.74762', '0.74762', '0.76205', '0.77086', '0.77067']
=========================
testing maxdeviation=5.2
Macro F1 score on test data: 0.77067
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '16.00', '35.00', '43.00', '45.00', '45.00', '47.00', '47.00', '50.00', '52.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.37643', '0.62136', '0.69468', '0.70196', '0.74762', '0.74762', '0.76205', '0.77086', '0.77067']
=========================
testing maxdeviation=5.3
Macro F1 score on test data: 0.77067
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '15.00', '35.00', '43.00', '45.00', '45.00', '47.00', '47.00', '50.00', '52.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.39893', '0.62136', '0.69468', '0.70196', '0.74762', '0.74762', '0.76205', '0.77086', '0.77067']
=========================
testing maxdeviation=5.4
Macro F1 score on test data: 0.77067
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '15.00', '35.00', '43.00', '45.00', '45.00', '47.00', '47.00', '50.00', '52.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.39893', '0.62136', '0.69468',

'0.70196', '0.74762', '0.74762', '0.76205', '0.77086', '0.77067']
=========================
testing maxdeviation=5.5
Macro F1 score on test data: 0.77067
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '15.00', '35.00', '43.00', '45.00', '45.00', '47.00', '47.00', '50.00', '52.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.39893', '0.62136', '0.69468', '0.70196', '0.74762', '0.74762', '0.76205', '0.77086', '0.77067']
=========================
testing maxdeviation=5.6
Macro F1 score on test data: 0.77067
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '14.00', '35.00', '43.00', '45.00', '45.00', '47.00', '47.00', '50.00', '52.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.39893', '0.62136', '0.69468', '0.70196', '0.74762', '0.74762', '0.76205', '0.77086', '0.77067']
=========================
testing maxdeviation=5.7
Macro F1 score on test data: 0.77067
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '9.00', '14.00', '35.00', '43.00', '45.00', '45.00', '47.00', '47.00', '50.00', '52.00']
average F1 score on test dataset: ['0.32310', '0.30251', '0.39893', '0.62136', '0.69468', '0.70196', '0.74762', '0.74762', '0.76205', '0.77086', '0.77067']
=========================
testing maxdeviation=5.8
Macro F1 score on test data: 0.77067
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '8.00', '14.00', '35.00', '43.00', '45.00', '45.00', '47.00', '47.00', '50.00', '52.00']
average F1 score on test dataset: ['0.32310', '0.29735', '0.39893', '0.62136', '0.69468', '0.70196', '0.74762', '0.74762', '0.76205', '0.77086', '0.77067']
=========================
testing maxdeviation=5.9
Macro F1 score on test data: 0.77067
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '8.00', '14.00', '35.00', '43.00', '45.00', '45.00', '47.00', '47.00', '50.00', '52.00']
average F1 score on test dataset: ['0.32310', '0.29735', '0.39893', '0.62136', '0.69468', '0.70196', '0.74762', '0.74762', '0.76205', '0.77086', '0.77067']
=========================
testing maxdeviation=6.0
Macro F1 score on test data: 0.77067

train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '8.00', '14.00', '35.00', '43.00', '45.00', '45.00', '47.00', '47.00', '50.00', '52.00']
average F1 score on test dataset: ['0.32310', '0.29735', '0.39893', '0.62136', '0.69468', '0.70196', '0.74762', '0.74762', '0.76205', '0.77086', '0.77067']
========================
testing maxdeviation=6.1
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '8.00', '14.00', '35.00', '42.00', '44.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.29735', '0.39893', '0.62136', '0.68526', '0.69259', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=6.2
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '8.00', '13.00', '35.00', '42.00', '44.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.29735', '0.41895', '0.62136', '0.68526', '0.69259', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=6.3
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '8.00', '13.00', '35.00', '42.00', '44.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.29735', '0.41895', '0.62136', '0.68526', '0.69259', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=6.4
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '8.00', '13.00', '35.00', '42.00', '44.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.29735', '0.41895', '0.62136', '0.68526', '0.69259', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=6.5
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '8.00', '13.00', '35.00', '42.00', '44.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.29735', '0.41895', '0.62136', '0.68526',

'0.69259', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=6.6
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '8.00', '13.00', '35.00', '42.00', '44.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.29735', '0.41895', '0.62136', '0.68526', '0.69259', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=6.7
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '7.00', '13.00', '34.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.29098', '0.41895', '0.63227', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=6.8
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '7.00', '13.00', '34.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.29098', '0.41895', '0.63227', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=6.9
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '7.00', '13.00', '34.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.29098', '0.41895', '0.63227', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=7.0
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '7.00', '13.00', '34.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.29098', '0.41895', '0.63227', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=7.1
Macro F1 score on test data: 0.76064

train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '6.00', '12.00', '34.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.28313', '0.37883', '0.63227', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=7.2
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '6.00', '12.00', '34.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.28313', '0.37883', '0.63227', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=7.3
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '6.00', '12.00', '34.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.28313', '0.37883', '0.63227', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=7.4
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '6.00', '12.00', '34.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.28313', '0.37883', '0.63227', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=7.5
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '6.00', '12.00', '34.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.28313', '0.37883', '0.63227', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=7.6
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '6.00', '12.00', '34.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.28313', '0.37883', '0.63227', '0.69385',

'0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
=========================
testing maxdeviation=7.7
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '6.00', '12.00', '34.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.28313', '0.37883', '0.63227', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
=========================
testing maxdeviation=7.8
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '6.00', '12.00', '34.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.28313', '0.37883', '0.63227', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
=========================
testing maxdeviation=7.9
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '6.00', '12.00', '34.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.28313', '0.37883', '0.63227', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
=========================
testing maxdeviation=8.0
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '6.00', '12.00', '34.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.28313', '0.37883', '0.63227', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
=========================
testing maxdeviation=8.1
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '6.00', '12.00', '34.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.28313', '0.37883', '0.63227', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
=========================
testing maxdeviation=8.2
Macro F1 score on test data: 0.76064

train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['8.00', '6.00', '12.00', '34.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.32310', '0.28313', '0.37883', '0.63227', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=8.3
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['7.00', '6.00', '12.00', '34.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.36804', '0.28313', '0.37883', '0.63227', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=8.4
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['7.00', '6.00', '12.00', '33.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.36804', '0.28313', '0.37883', '0.61727', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=8.5
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['7.00', '6.00', '12.00', '33.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.36804', '0.28313', '0.37883', '0.61727', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=8.6
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['7.00', '6.00', '12.00', '33.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.36804', '0.28313', '0.37883', '0.61727', '0.69385', '0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=8.7
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['7.00', '4.00', '12.00', '33.00', '41.00', '43.00', '44.00', '46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.36804', '0.23383', '0.37883', '0.61727', '0.69385',

'0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=8.8
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['7.00', '4.00', '12.00', '33.00', '41.00', '43.00', '44.00',
'46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.36804', '0.23383', '0.37883', '0.61727', '0.69385',
'0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=8.9
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['7.00', '4.00', '12.00', '33.00', '41.00', '43.00', '44.00',
'46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.36804', '0.23383', '0.37883', '0.61727', '0.69385',
'0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=9.0
Macro F1 score on test data: 0.76064
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['7.00', '4.00', '12.00', '33.00', '41.00', '43.00', '44.00',
'46.00', '46.00', '49.00', '51.00']
average F1 score on test dataset: ['0.36804', '0.23383', '0.37883', '0.61727', '0.69385',
'0.70092', '0.73760', '0.73760', '0.75199', '0.76110', '0.76064']
========================
testing maxdeviation=9.1
Macro F1 score on test data: 0.75659
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['7.00', '4.00', '12.00', '33.00', '39.00', '41.00', '42.00',
'44.00', '44.00', '47.00', '49.00']
average F1 score on test dataset: ['0.36804', '0.23383', '0.37883', '0.61727', '0.69005',
'0.69713', '0.73355', '0.73355', '0.74793', '0.75715', '0.75659']
========================
testing maxdeviation=9.2
Macro F1 score on test data: 0.75659
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['7.00', '4.00', '12.00', '33.00', '39.00', '41.00', '42.00',
'44.00', '44.00', '47.00', '49.00']
average F1 score on test dataset: ['0.36804', '0.23383', '0.37883', '0.61727', '0.69005',
'0.69713', '0.73355', '0.73355', '0.74793', '0.75715', '0.75659']
========================
testing maxdeviation=9.3
Macro F1 score on test data: 0.75659

train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['7.00', '4.00', '12.00', '33.00', '39.00', '41.00', '42.00', '44.00', '44.00', '47.00', '49.00']
average F1 score on test dataset: ['0.36804', '0.23383', '0.37883', '0.61727', '0.69005', '0.69713', '0.73355', '0.73355', '0.74793', '0.75715', '0.75659']
==========================
testing maxdeviation=9.4
Macro F1 score on test data: 0.75659
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['7.00', '4.00', '12.00', '33.00', '39.00', '41.00', '42.00', '44.00', '44.00', '47.00', '49.00']
average F1 score on test dataset: ['0.36804', '0.23383', '0.37883', '0.61727', '0.69005', '0.69713', '0.73355', '0.73355', '0.74793', '0.75715', '0.75659']
==========================
testing maxdeviation=9.5
Macro F1 score on test data: 0.75659
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['7.00', '4.00', '12.00', '33.00', '39.00', '41.00', '42.00', '44.00', '44.00', '47.00', '49.00']
average F1 score on test dataset: ['0.36804', '0.23383', '0.37883', '0.61727', '0.69005', '0.69713', '0.73355', '0.73355', '0.74793', '0.75715', '0.75659']
==========================
testing maxdeviation=9.6
Macro F1 score on test data: 0.75659
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['7.00', '4.00', '12.00', '33.00', '39.00', '41.00', '42.00', '44.00', '44.00', '47.00', '49.00']
average F1 score on test dataset: ['0.36804', '0.23383', '0.37883', '0.61727', '0.69005', '0.69713', '0.73355', '0.73355', '0.74793', '0.75715', '0.75659']
==========================
testing maxdeviation=9.7
Macro F1 score on test data: 0.75659
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['7.00', '4.00', '12.00', '33.00', '39.00', '41.00', '42.00', '44.00', '44.00', '47.00', '49.00']
average F1 score on test dataset: ['0.36804', '0.23383', '0.37883', '0.61727', '0.69005', '0.69713', '0.73355', '0.73355', '0.74793', '0.75715', '0.75659']
==========================
testing maxdeviation=9.8
Macro F1 score on test data: 0.75659
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['6.00', '4.00', '12.00', '33.00', '39.00', '41.00', '42.00', '44.00', '44.00', '47.00', '49.00']
average F1 score on test dataset: ['0.31966', '0.23383', '0.37883', '0.61727', '0.69005',

'0.69713', '0.73355', '0.73355', '0.74793', '0.75715', '0.75659']
=========================
testing maxdeviation=9.9
Macro F1 score on test data: 0.75659
train dataset size: [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1038]
average node count on test dataset: ['6.00', '4.00', '12.00', '33.00', '39.00', '41.00', '42.00', '44.00', '44.00', '47.00', '49.00']
average F1 score on test dataset: ['0.31966', '0.23383', '0.37883', '0.61727', '0.69005', '0.69713', '0.73355', '0.73355', '0.74793', '0.75715', '0.75659']
=========================