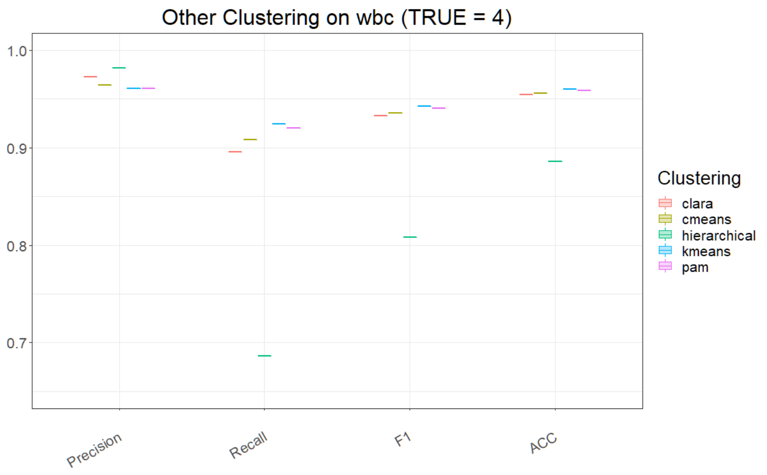
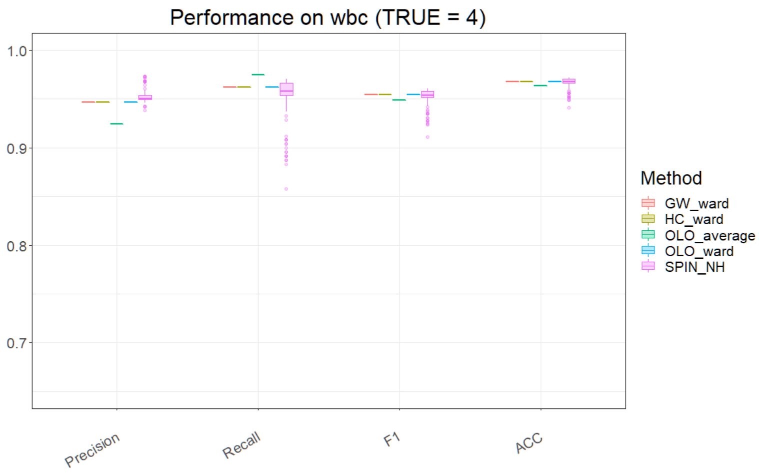
WBC：

WBC Dataset			
	Atribute ID	Types	Value Range
Atrributes	2	Clump Thickness	1 -10
	3	Uniformity of Cell Size	1 -10
	4	Uniformity of Cell Shape	1 -10
	5	Marginal Adhesion	1 -10
	6	Single Epithelial Cell Size	1 -10
	7	Bare Nuclei	1 -10
	8	Bland Chromatin	1 -10
	9	Normal Nucleoli	1 -10
	10	Mitoses	1 -10
Class Distribution	11	**2 = benign**	**4 = malignant**
		458 (65.5522%)	241 (34.4778%)
Number of Instances	699		
Number of Attributes	11 (ID, diagnosis, 9 real-valued input features)		
Missing attribute values	16		

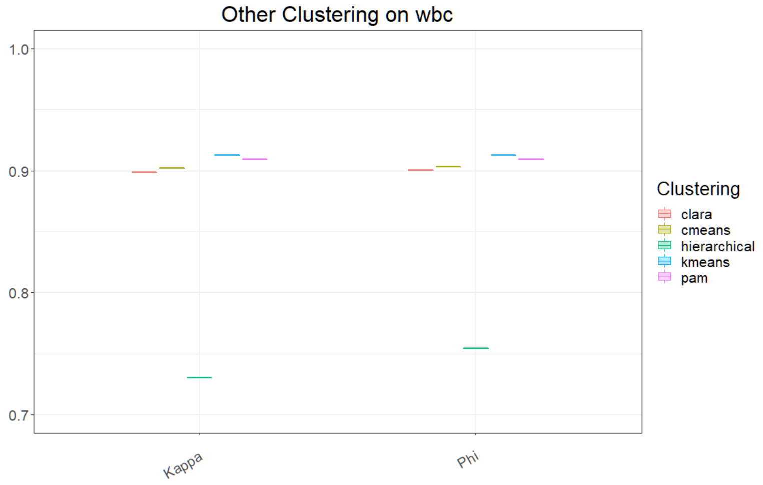
Since there are 16 missing instances in WBC, these samples have been removed and only analyzed the remaining 683 samples.

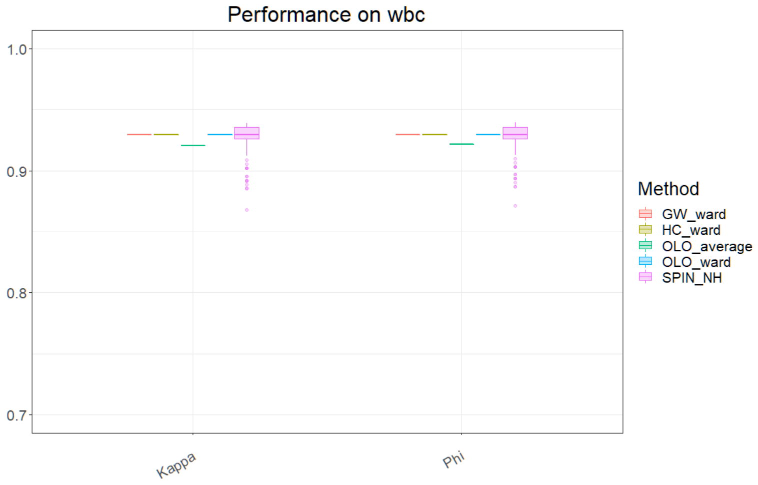
Exact Clustering:


Other Clustering on wbc (TRUE = 4)


Performance on wbc (TRUE = 4)

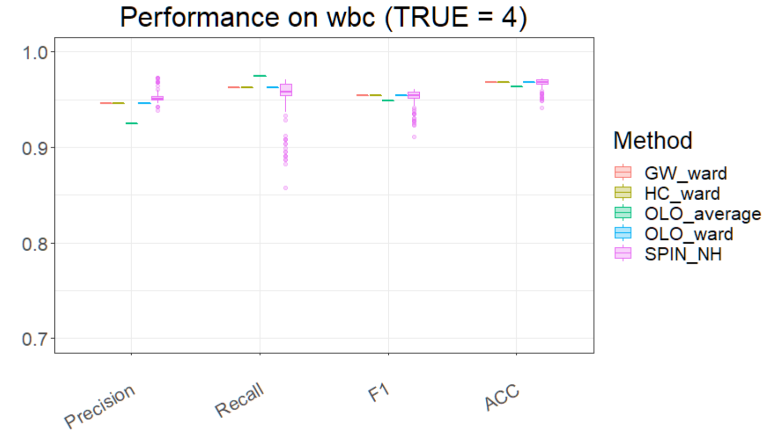
First of all, we can find that both the commonly used clustering methods (K-Means, C-Means, Hierarchical, Clara and Pam) and the Czeknowski's clustering perform well on the dataset WBC. With the exception of Hierarchical, the other algorithms can achieve an accuracy of around 95%.
It can be seen from the figure above that when diagnosing whether a tumor is malignant, although the precision of other clustering algorithms is higher than that of Czeknowski's clustering algorithm, the accuracy, F1 score and recall of our algorithm are better than others, especially recall. Therefore, I believe our algorithm is more sensitive in diagnosing malignant tumors and is more suitable as a judgment method for early diagnosis.

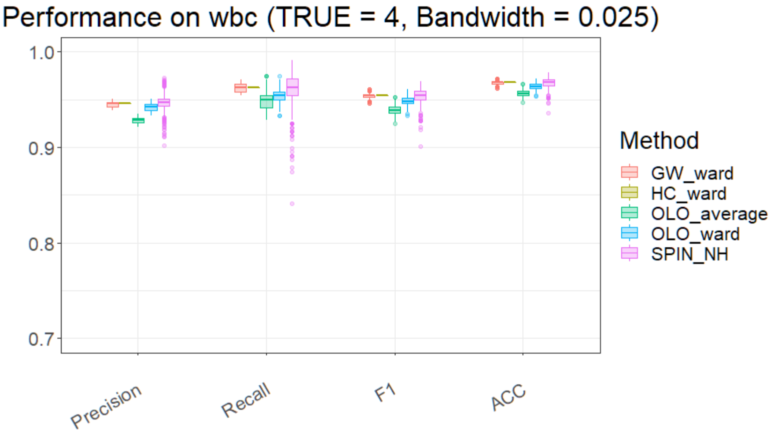


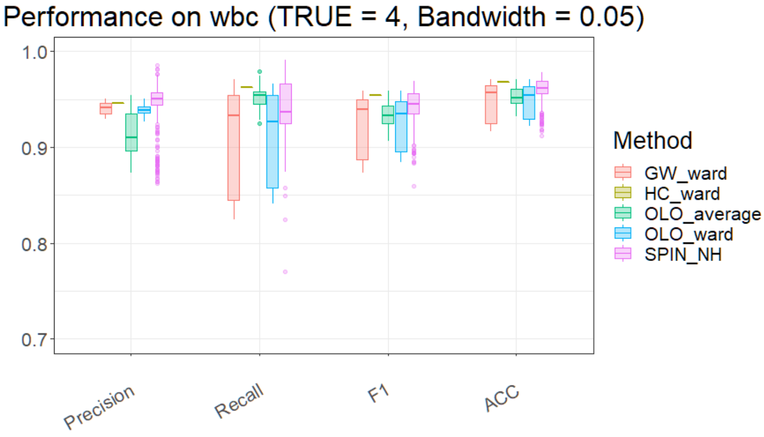


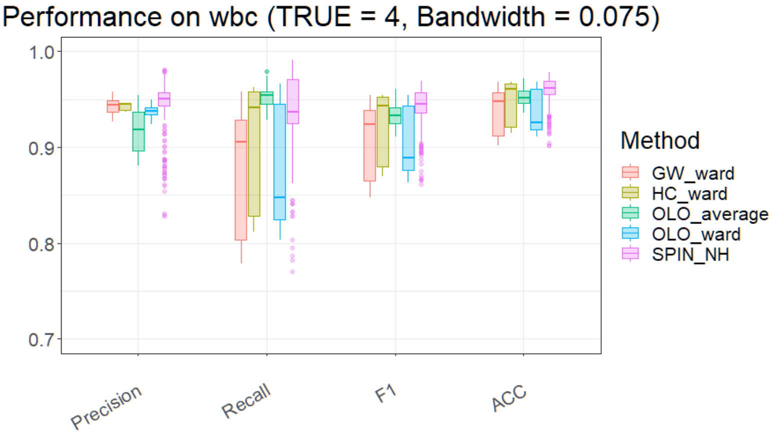
Because the proportion of TRUE and FALSE samples of the data is not equal, I also tried to compare and evaluate the algorithm through other measures. The above plots show the results of the Kappa statistic and Phi coefficient for our algorithm and its comparison algorithms. It can be seen that the two index results are almost identical, and the two metrics of our algorithm are better than those of the control algorithm.

Fuzzy Clustering:









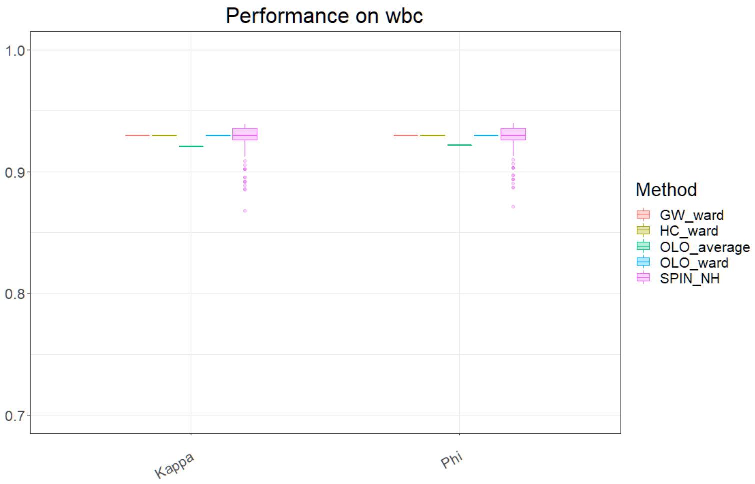
Accuracy on WBC

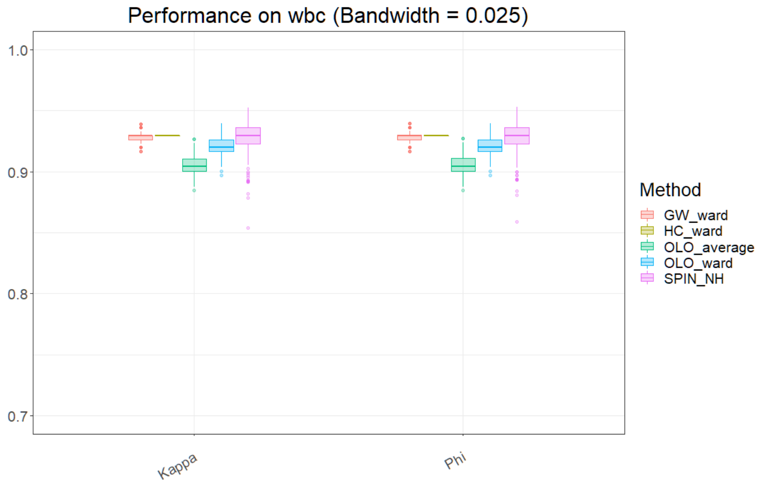
	Mean				Median				Max				Min			
	B=0	B=0.025	B=0.05	B=0.075	B=0	B=0.025	B=0.05	B=0.075	B=0	B=0.025	B=0.05	B=0.075	B=0	B=0.025	B=0.05	B=0.075
SPIN_NH	0.9678	0.9671	0.9603	0.9602	0.9678	0.9678	0.9619	0.9619	0.9722	0.978	0.978	0.978	0.9414	0.9356	0.9122	0.9019
OLO_ward	0.9678	0.9636	0.9466	0.939	0.9678	0.9634	0.9546	0.9261	0.9678	0.9722	0.9707	0.9678	0.9678	0.9531	0.9224	0.9107
GW_ward	0.9678	0.9672	0.945	0.9358	0.9678	0.9678	0.9575	0.948	0.9678	0.9722	0.9707	0.9678	0.9678	0.9619	0.9165	0.9019
HC_ward	0.9678	0.9678	0.9678	0.9471	0.9678	0.9678	0.9678	0.9605	0.9678	0.9678	0.9678	0.9678	0.9678	0.9678	0.9678	0.9151
OLO_average	0.9634	0.9568	0.9526	0.9525	0.9634	0.9561	0.9517	0.9517	0.9634	0.9663	0.9707	0.9722	0.9634	0.9473	0.9327	0.9356

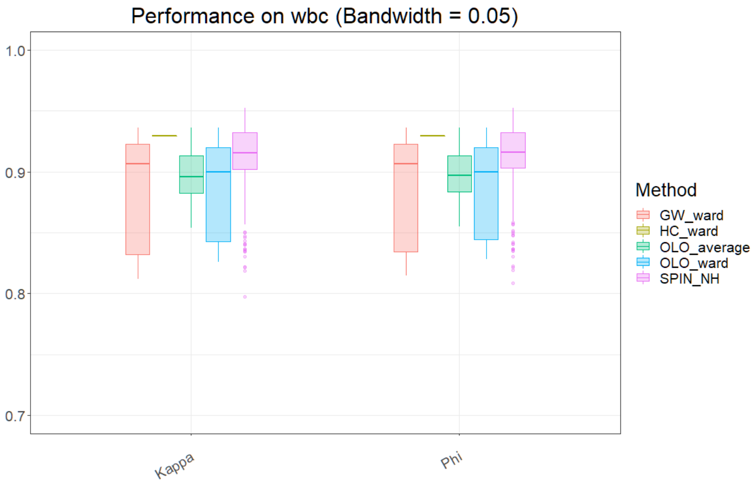
It can be seen that for WBC, moderately widening the bandwidth cannot improve the algorithm's overall performance. I believe the reason for this is that fuzzy Czekanowski's clustering is a compromise algorithm between exact Czekanowski's clustering and random sampling weighted by fuzzy C-Means membership values. Since C-Means does not perform better than exact Czekanowski's clustering, increasing the bandwidth of fuzzy clustering will not improve performance either.

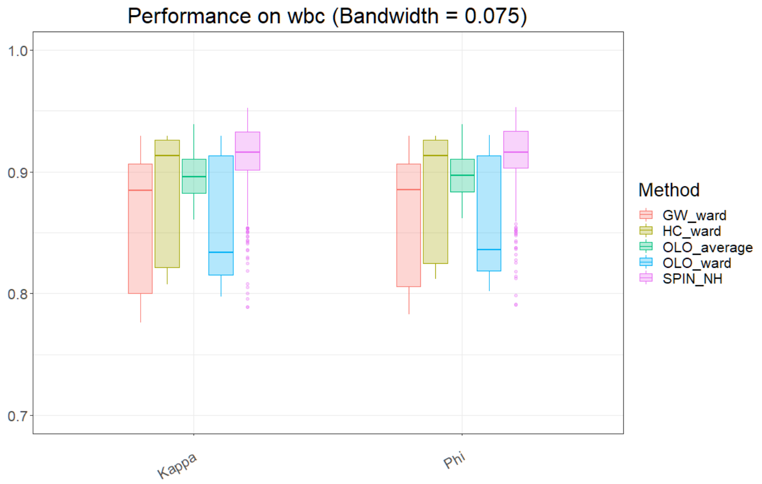
A wider bandwidth will lead to instability of our clustering algorithm. However, for most serialization methods, wider bandwidth also allows the algorithm to find better performance.

In particular, the serialization method SPIN_NH (neighborhood algorithm) has results that are inconsistent over multiple attempts. This means that even with exact clustering, the clustering results can vary. With increased bandwidth, the use of SPIN_NH has the opportunity to give the best clustering results.


Performance on wbc


Performance on wbc (Bandwidth = 0.025)


Performance on wbc (Bandwidth = 0.05)


Performance on wbc (Bandwidth = 0.075)

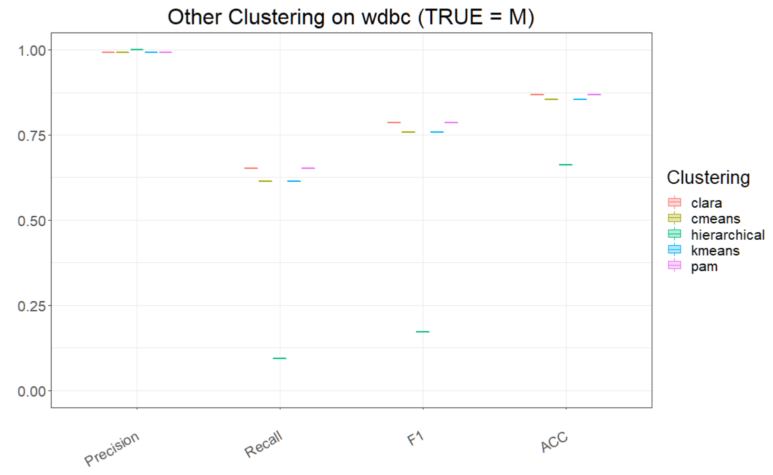
Like the previous conclusion, the Kappa statistic and Phi coefficient results are almost the same.

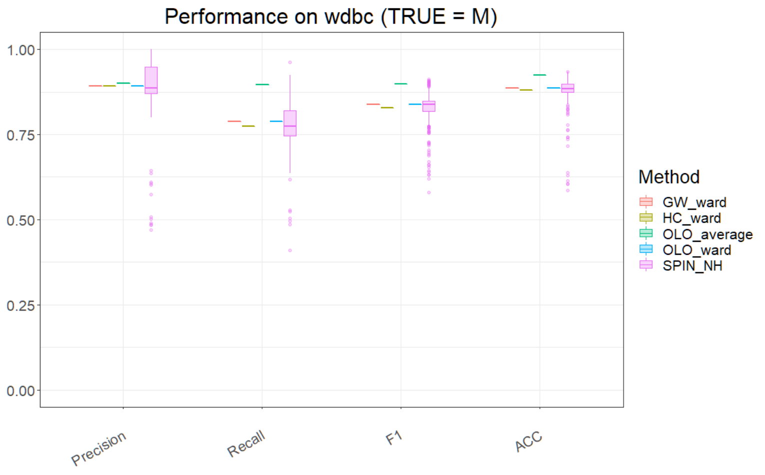
Wider bandwidths can lead to erratic clustering performance, but it is also possible to appropriately improve the best performance of the algorithm.

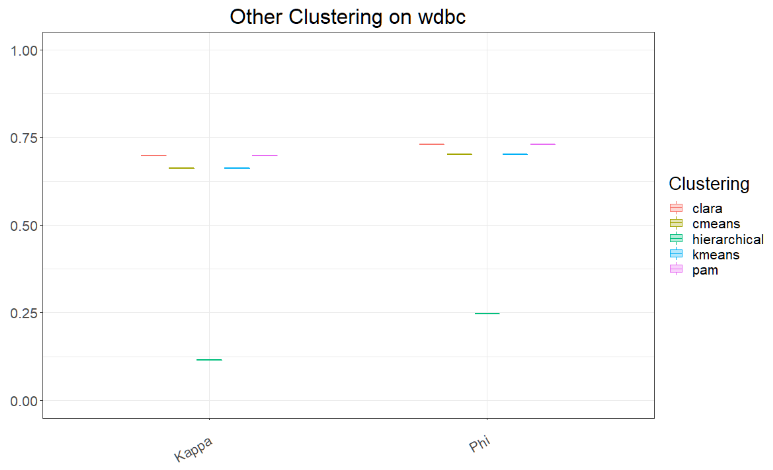
WDBC:

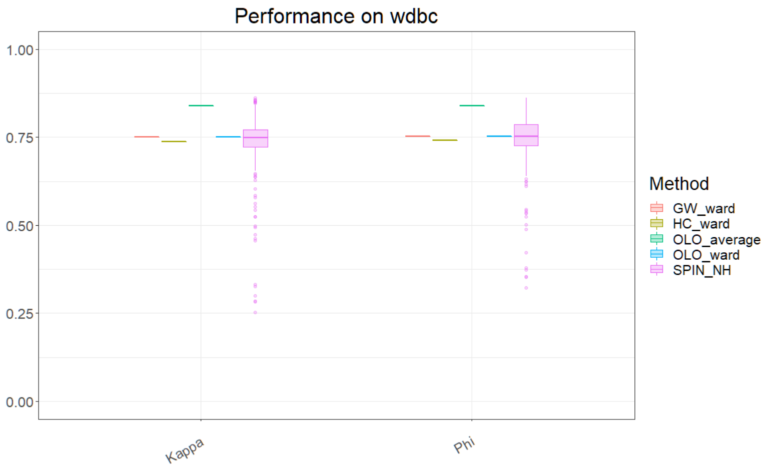
WDBC Dataset					
	Atribute ID	Types	Value Range		
			Mean	Standard Error	Worst/Largest
Atrributes	3, 13, 23	Radius	6.981 - 28.110	0.112 - 2.873	7.930 - 36.040
	4, 14, 24	Texture	9.710 - 39.280	0.360 - 4.885	12.020 - 49.540
	5, 15, 25	Perimeter	43.790 - 188.500	0.757 - 21.980	50.410 - 251.200
	6, 16, 26	Area	143.500 - 2501.000	6.802 - 542.200	185.200 - 4254.000
	7, 17, 27	Smoothness	0.053 - 0.163	0.0017 - 0.0311	0.071 - 0.223
	8, 18, 28	Compactness	0.019 - 0.345	0.002 - 0.135	0.027 - 1.058
	9, 19, 29	Concavity	0.000 - 0.427	0.000 - 0.396	0.000 - 1.252
	10, 20, 30	Concave Points	0.000 - 0.201	0.000 - 0.053	0.000 - 0.291
	11, 21, 31	Symmetry	0.106 - 0.304	0.008 - 0.079	0.157 - 0.664
	12, 22, 32	Fractal Dimension	0.050 - 0.097	0.001- 0.030	0.055 - 0.208
Class Distribution	2	B = benign		M = malignant	
		357 (62.7417%)		212 (37.2584%)	
Number of Instances	569				
Number of Attributes	32 (ID, diagnosis, 30 real-valued input features)				
Missing attribute values	None				

Exact Clustering:







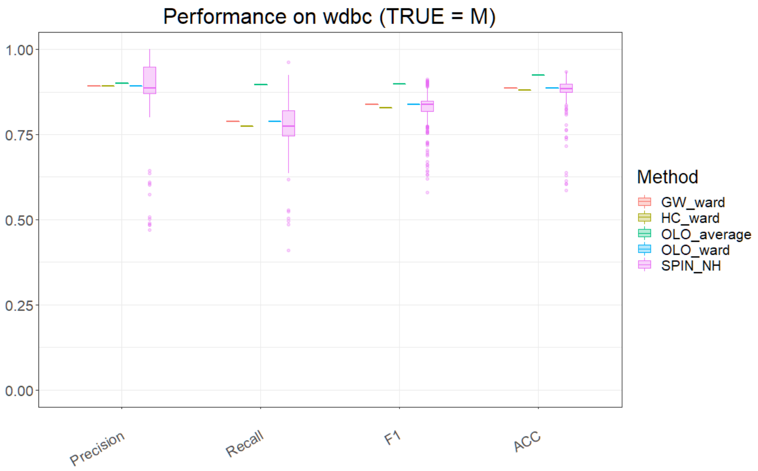


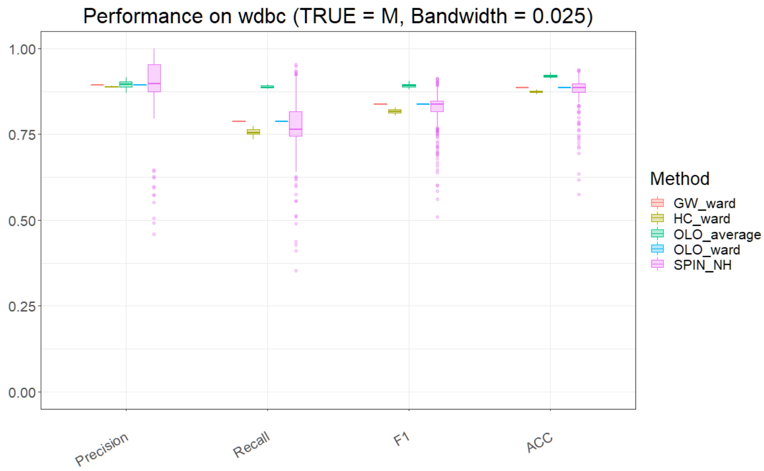
It can be noticed that, as with the WBC dataset, the precision results for the control group are better than those of our clustering algorithm. Especially for the hierarchical, its precision even reaches 100%.

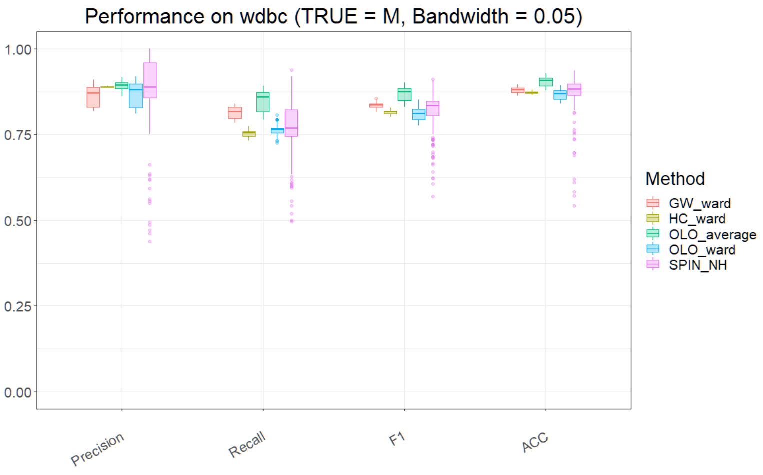
However, in addition to precision, Czeknowski's clustering outperforms these five common clustering algorithms in terms of recall, F1 score, accuracy, Kappa statistics and Phi coefficient. In particular, the hierarchical clustering algorithm, which performs exceptionally well on precision, has significantly lower results on these metrics than other algorithms.

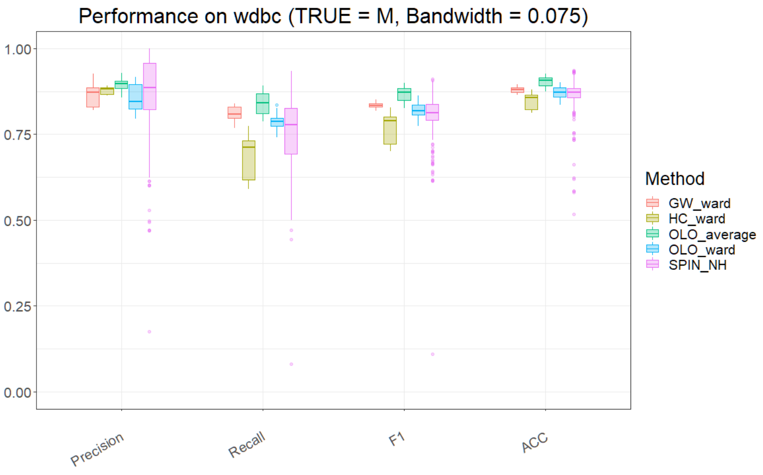
Additionally, OLO_average shows a clear advantage over other serialisation methods.

Fuzzy Clustering:


Performance on wdbc (TRUE = M)


Performance on wdbc (TRUE = M, Bandwidth = 0.025)

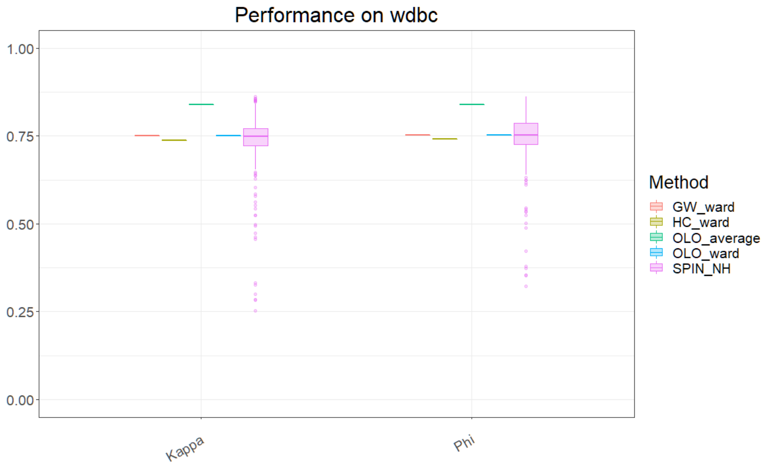

Performance on wdbc (TRUE = M, Bandwidth = 0.05)


Performance on wdbc (TRUE = M, Bandwidth = 0.075)

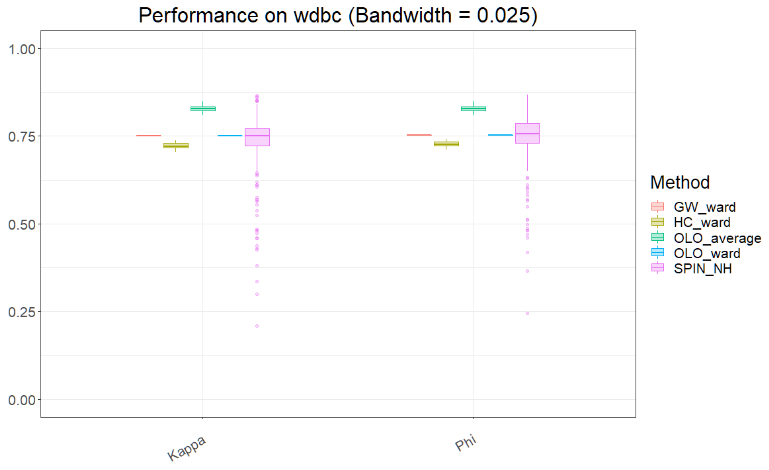
Accuracy on WDBC

	Mean				Median				Max				Min			
	B=0	B=0.025	B= 0.05	B=0.075	B=0	B=0.025	B=0.05	B=0.075	B=0	B=0.025	B=0.05	B=0.075	B=0	B=0.025	B=0.05	B=0.075
SPIN_NH	0.8800	0.8810	0.8751	0.8665	0.8840	0.8866	0.8822	0.8717	0.9350	0.9385	0.9367	0.9367	0.5852	0.5747	0.5413	0.5167
OLO_ward	0.8858	0.8858	0.8663	0.8715	0.8858	0.8858	0.8682	0.8717	0.8858	0.8858	0.8946	0.9016	0.8858	0.8858	0.8401	0.8366
GW_ward	0.8858	0.8858	0.8799	0.8797	0.8858	0.8858	0.8805	0.8805	0.8858	0.8858	0.8963	0.8963	0.8858	0.8858	0.8629	0.8647
HC_ward	0.8805	0.8740	0.8734	0.8461	0.8805	0.8735	0.8735	0.8576	0.8805	0.8805	0.8805	0.8805	0.8805	0.8664	0.8647	0.8120
OLO_average	0.9244	0.9199	0.9044	0.9042	0.9244	0.9192	0.9077	0.9069	0.9244	0.9297	0.9279	0.9262	0.9244	0.9104	0.8787	0.8752

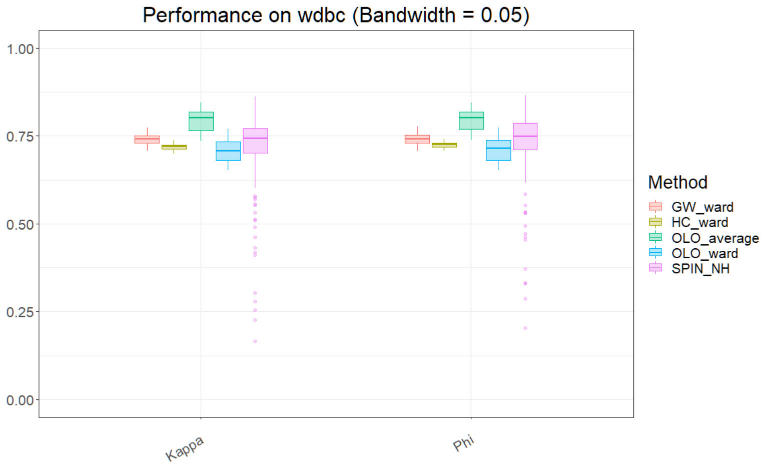
For the WDBC dataset, the gradual widening of the bandwidth from 0 to 0.075 does not show a clear tendency towards instability as it does for WBC.



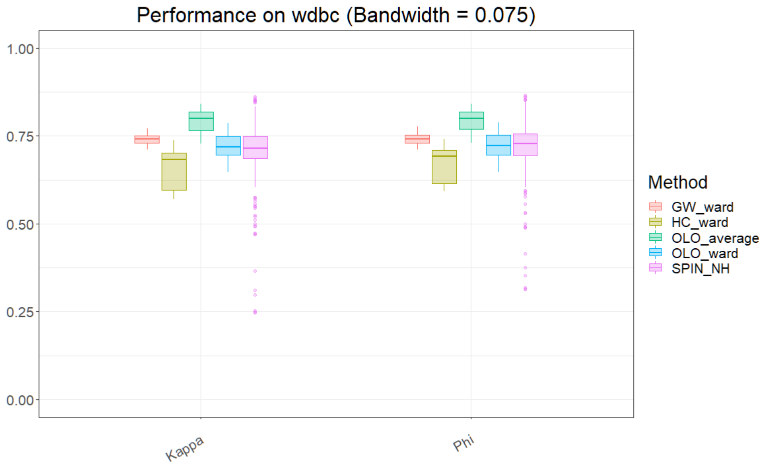
Performance on wdbc



Performance on wdbc (Bandwidth = 0.025)



Performance on wdbc (Bandwidth = 0.05)



Performance on wdbc (Bandwidth = 0.075)

Kappa & Phi on WDBC																
	Kappa								Phi							
	Max				Min				Max				Min			
	B = 0	B = 0.025	B = 0.05	B = 0.075	B = 0	B = 0.025	B = 0.05	B = 0.075	B = 0	B = 0.025	B = 0.05	B = 0.075	B = 0	B = 0.025	B = 0.05	B = 0.075
SPIN_NH	0.8610	0.8660	0.8620	0.8618	0.2531	0.2099	0.1652	-0.1617	0.8611	0.8683	0.8646	0.8649	0.3227	0.2450	0.2038	-0.1850
OLO_ward	0.7497	0.7497	0.7696	0.7858	0.7497	0.7497	0.6516	0.6467	0.7532	0.7532	0.7724	0.7877	0.7532	0.7532	0.6533	0.6473
GW_ward	0.7497	0.7497	0.7728	0.7710	0.7497	0.7497	0.7062	0.7103	0.7532	0.7532	0.7764	0.7774	0.7532	0.7532	0.7063	0.7103
HC_ward	0.7373	0.7373	0.7373	0.7373	0.7373	0.7041	0.6999	0.5702	0.7417	0.7417	0.7417	0.7417	0.7417	0.7111	0.7073	0.5920
OLO_average	0.8382	0.8491	0.8451	0.8412	0.8382	0.8085	0.7358	0.7287	0.8382	0.8492	0.8453	0.8414	0.8382	0.8085	0.7378	0.7302

By analyzing the results of Kappa statistic and Phi coefficient, we can clearly see the impact of bandwidth on the clustering results. When the bandwidth reaches 0.075, the minimum results of Kappa statistic and Phi coefficient are even lower than 0.

WPBC:

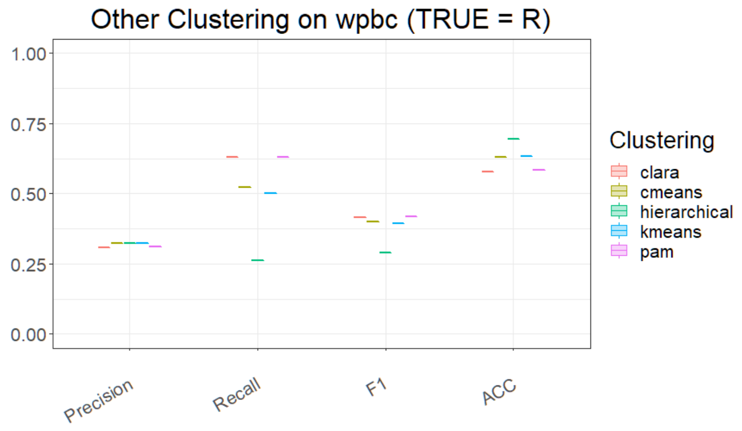
	Atribute ID	Types	Value Range		
			Mean	Standard Error	Worst/Largest
Atrributes	3	Time	46.94	2.479	1.000 - 125.000
	4, 14, 24	Radius	10.950 - 27.220	0.194 - 1.819	12.840 - 35.130
	5, 15, 25	Texture	10.380 - 39.280	0.362 - 3.503	16.670 - 49.540
	6, 16, 26	Perimeter	71.900 - 182.100	1.153 - 13.280	85.100 - 232.200
	7, 17, 27	Area	361.600 - 2250.000	13.990 - 316.000	508.100 - 3903.000
	8, 18, 28	Smoothness	0.075 - 0.145	0.003 - 0.031	0.082 - 0.223
	9, 19, 29	Compactness	0.046 - 0.311	0.007 - 0.135	0.051 - 1.058
	10, 20, 30	Concavity	0.024 - 0.427	0.011 - 0.144	0.024 - 1.170
	11, 21, 31	Concave Points	0.020 - 0.201	0.005 - 0.039	0.029 - 0.290
	12, 22, 32	Symmetry	0.131 - 0.304	0.008 - 0.060	0.157 - 0.664
	13, 23, 33	Fractal Dimension	0.0503 - 0.097	0.001 - 0.013	0.055 - 0.208
	34	Tumor Size	2.868	0.14	0.400 - 10.000
	35	Lymph node status	3.211	0.393	0.000 - 27.000
Class Distribution	2	N = nonrecur		R = recur	
		151 (76.2626%)		47 (23.7374%)	
Number of Instances	198				
Number of Attributes	35 (ID, outcome, time, 32 real-valued input features)				
Missing attribute values	4				

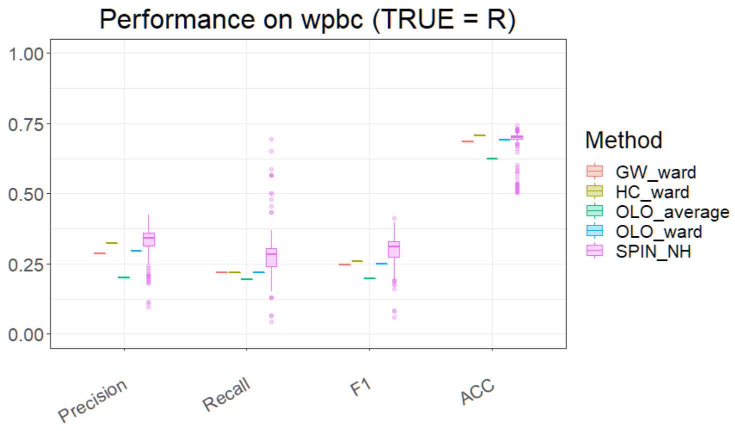
Since there are 4 missing instances in WPBC, these samples have been removed and only analyzed the remaining 194 samples.

The feature Time records the time of two states (recurrence time if field 2 = R; disease-free time if field 2 = N), so it will not be analyzed.

The analysis characteristics are feature 4 to feature 35.

Exact Clustering:



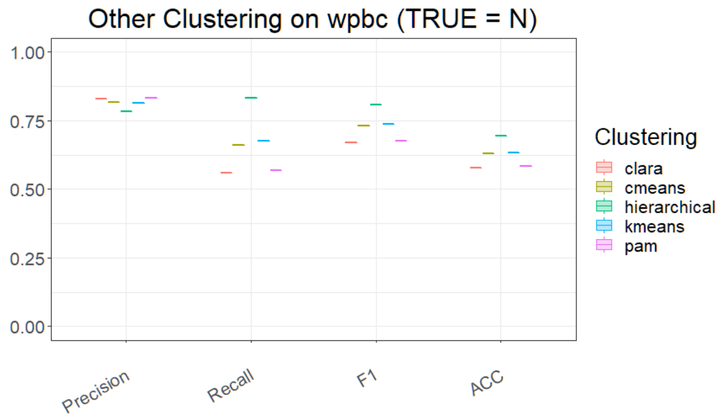


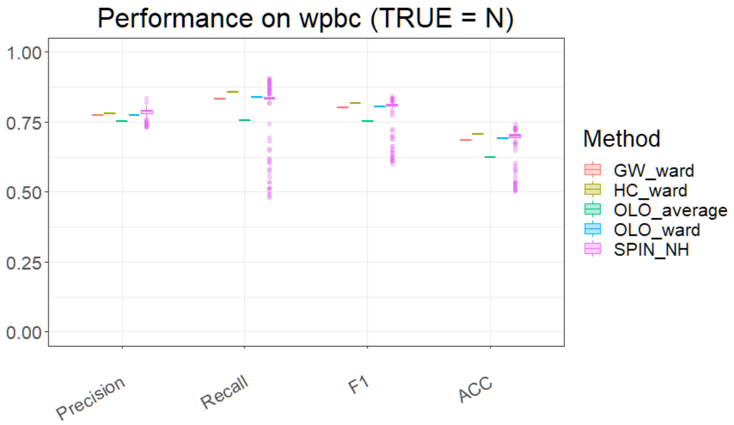
Unlike the datasets WBC and WDBC, none of the clustering methods attempted so far has achieved satisfactory results on WPBC. No of them has been able to exceed 75% accuracy.

If recur is used as the true label, the precision values only reach about 30%. Recall and F1 were also unsatisfactory.

The Czekanowski's clustering only slightly exceeded the control group in terms of precision.

Since the results using recur as true were not ideal, I also analyzed the case of setting ‘N’ (nonrecur) as true.

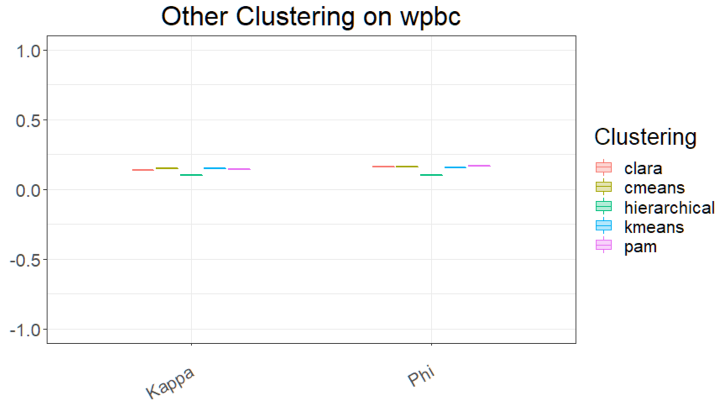

Other Clustering on wpbc (TRUE = N)

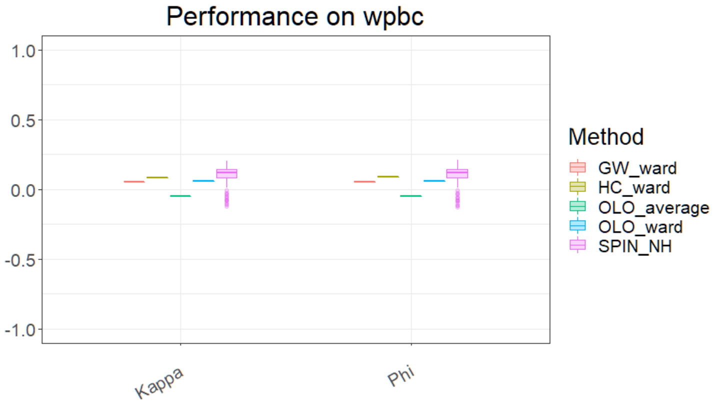

Performance on wpbc (TRUE = N)

When using nonrecur as the true label, it can be found that the results of precision, recall and F1 are significantly improved. The results of Czekanowski's clustering method are overall better than the control methods regarding recall, F1 score and accuracy.

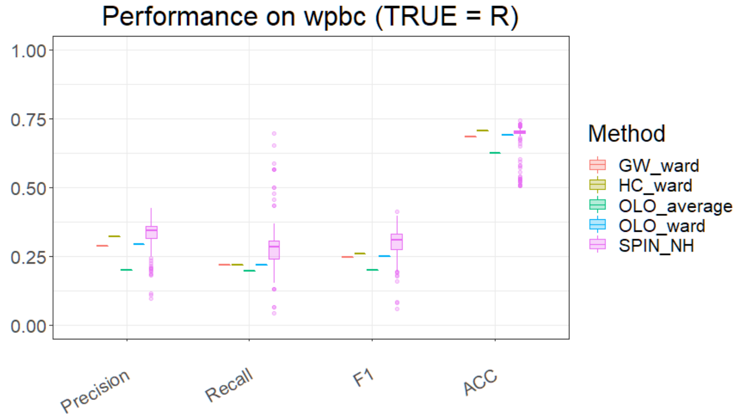
Unlike the performance on WDBC, OLO_average is the worst performer on WPBC.

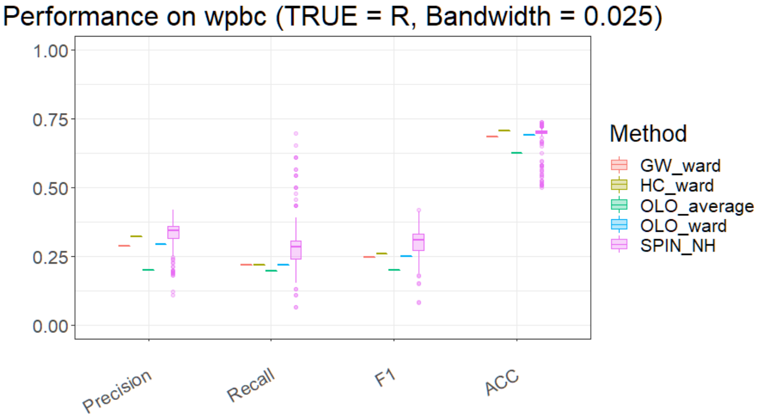
Therefore, I believe that WPBC is not ideal for our paper to demonstrate the advantages of Czekanowski's clustering.

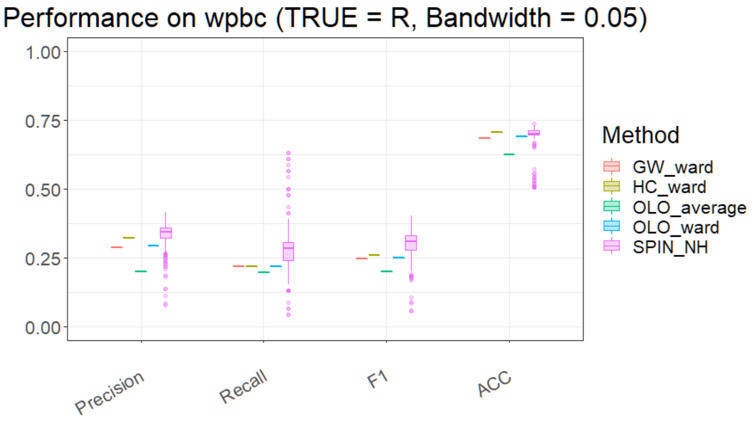

Other Clustering on wpbc

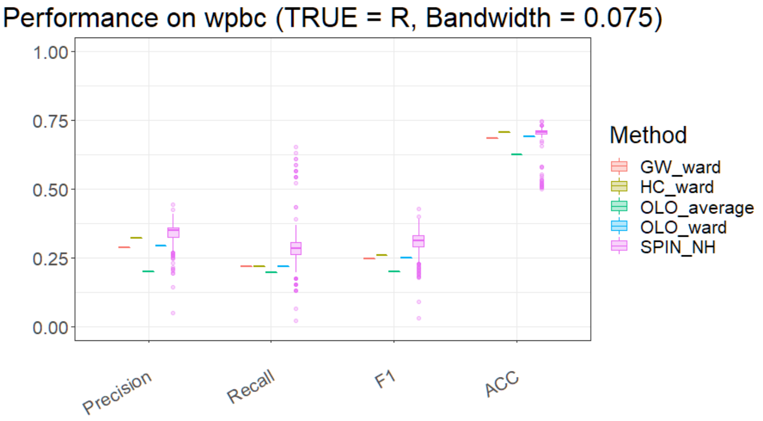

Performance on wpbc

Fuzzy Clustering:


Performance on wpbc (TRUE = R)

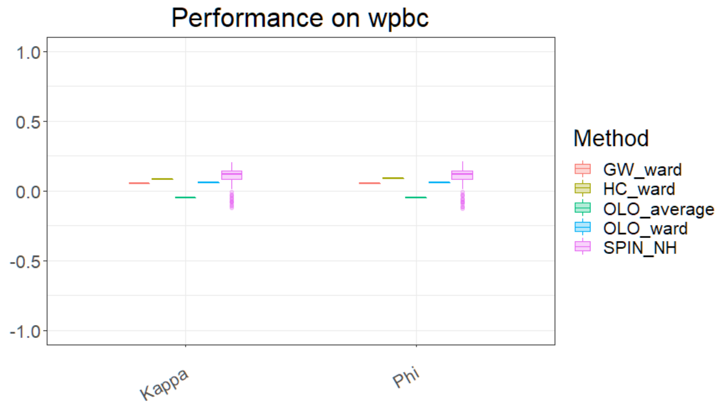

Performance on wpbc (TRUE = R, Bandwidth = 0.025)

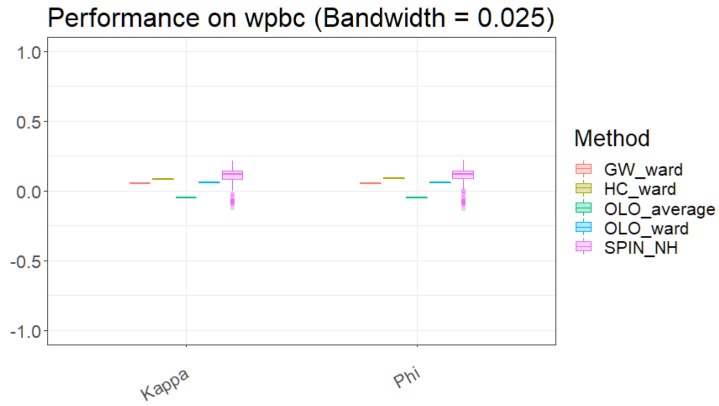

Performance on wpbc (TRUE = R, Bandwidth = 0.05)

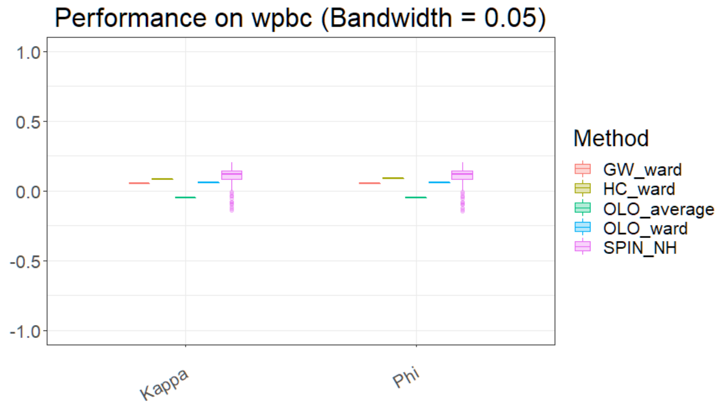

Performance on wpbc (TRUE = R, Bandwidth = 0.075)

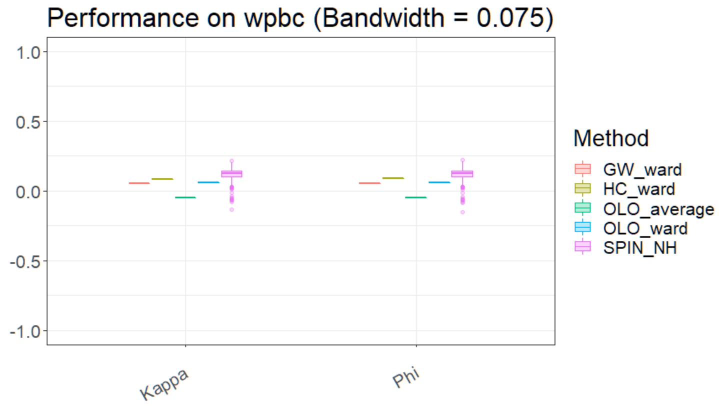
Accuracy on WPBC																
	Mean				Median				Max				Min			
	B=0	B=0.025	B= 0.05	B=0.075	B=0	B=0.025	B=0.05	B=0.075	B=0	B=0.025	B=0.05	B=0.075	B=0	B=0.025	B=0.05	B=0.075
SPIN_NH	0.6962	0.6952	0.6975	0.6970	0.7010	0.7010	0.7010	0.7062	0.7423	0.7371	0.7371	0.7474	0.5052	0.5000	0.5052	0.5000
OLO_ward	0.6907	0.6907	0.6907	0.6907	0.6907	0.6907	0.6907	0.6907	0.6907	0.6907	0.6907	0.6907	0.6907	0.6907	0.6907	0.6907
GW_ward	0.6856	0.6856	0.6856	0.6856	0.6856	0.6856	0.6856	0.6856	0.6856	0.6856	0.6856	0.6856	0.6856	0.6856	0.6856	0.6856
HC_ward	0.7062	0.7062	0.7062	0.7062	0.7062	0.7062	0.7062	0.7062	0.7062	0.7062	0.7062	0.7062	0.7062	0.7062	0.7062	0.7062
OLO_average	0.6237	0.6237	0.6237	0.6237	0.6237	0.6237	0.6237	0.6237	0.6237	0.6237	0.6237	0.6237	0.6237	0.6237	0.6237	0.6237

As can be seen from the results, increasing the bandwidth to 0.075 has no significant effect on the overall clustering performance.


Performance on wpbc


Performance on wpbc (Bandwidth = 0.025)


Performance on wpbc (Bandwidth = 0.05)


Performance on wpbc (Bandwidth = 0.075)

The previous conclusions can be verified by analyzing the Kappa statistics and Phi coefficients. Increasing the bandwidth does not improve the performance of the clustering.

The results of the Kappa statistics and Phi coefficients for OLO_average and some SPIN_NH are even below 0.