

1. Resultados CC

El análisis estadístico se hizo entre 21 modelos (poblaciones) con 100 muestras emparejadas.

El nivel de significancia para los test globales (*family-wise*) es de $\alpha = 0,050$. . No hemos podido rechazar la hipótesis nula de que la población sea normal para ninguna población (mínimo $p = 0,029$). Por lo tanto asumimos que las poblaciones son normales.

Hemos aplicado el test Bartlett para evaluar la homogeneidad de los datos y rechazamos la hipótesis nula ($p \approx 0$) de que los datos son homocedásticos. Por lo que asumimos que los datos son heteroscedásticos.

Dado que tenemos más de dos poblaciones, donde todas son normales pero heterocedásticas, usamos el test no paramétrico de Friedman como test *omnibus* para determinar si hay alguna diferencia significativa entre las medias de las poblaciones. Utilizamos el test de Nemenyi como *post-hoc* para saber que diferencias son significativas. Mostramos, en la tabla 1, la media (M), la desviación estándar (SD), y el ranking medio (MR) de cada modelo a partir de todas las muestras. Consideraremos significativas la diferencias entre modelos si la diferencia entre rankings medios es mayor que la distancia crítica de $CD = 3,132$ dada por el test de Nemenyi.

	MR	M	SD	CI	d	Magnitude
DUMMY	14.905	0.500	0.000	[0.500, 0.500]	-	large
PLS + KNNSScaler	13.565	0.500	0.052	[0.484, 0.517]	-0.013	negligible
PCA + KNNSScaler	13.385	0.503	0.051	[0.487, 0.519]	-0.078	negligible
PLS + SVCSScaler	13.345	0.504	0.054	[0.488, 0.521]	-0.116	negligible
PLS + KNN	13.155	0.505	0.057	[0.487, 0.523]	-0.122	negligible
mRMR + KNN	12.035	0.515	0.050	[0.500, 0.531]	-0.435	small
whole + SVC	11.905	0.519	0.048	[0.504, 0.534]	-0.573	medium
PLS + SVC	11.770	0.519	0.052	[0.502, 0.535]	-0.505	medium
mRMR + KNNSScaler	11.505	0.519	0.051	[0.503, 0.535]	-0.527	medium
PCA + SVC	11.030	0.525	0.055	[0.508, 0.542]	-0.649	medium
mRMR + SVC	10.785	0.529	0.054	[0.512, 0.545]	-0.752	medium
whole + SVCSScaler	10.785	0.526	0.041	[0.513, 0.539]	-0.908	large
mRMR + SVCSScaler	10.755	0.528	0.053	[0.512, 0.545]	-0.750	medium
whole + LR	10.700	0.527	0.038	[0.515, 0.539]	-0.993	large
PLS + LR	10.600	0.527	0.051	[0.511, 0.542]	-0.742	medium
PCA + SVCSScaler	10.400	0.525	0.054	[0.509, 0.542]	-0.668	medium
whole + KNN	9.355	0.538	0.048	[0.523, 0.553]	-1.128	large
whole + KNNSScaler	8.530	0.545	0.047	[0.530, 0.559]	-1.327	large
PCA + KNN	8.135	0.548	0.050	[0.532, 0.563]	-1.356	large
mRMR + LR	8.015	0.549	0.050	[0.533, 0.564]	-1.362	large
PCA + LR	6.340	0.566	0.061	[0.547, 0.585]	-1.520	large

Tabla 1: Comparativa del ROC AUC de los modelos con el dataset CC

Rechazamos la hipótesis nula ($p = 8,5 \times 10^{-38}$) del test de Friedman de que no hay diferencia entre las medidas de tendencia central de los modelos. Por lo que asumimos que hay una diferencia estadísticamente significativa entre las medianas de las poblaciones.

Basándonos en el test *post-hoc* de Nemenyi, asumimos que no hay diferencia significativa dentro de los siguientes grupos:

- DUMMY, PLS + KNNSScaler, PCA + KNNSScaler, PLS + SVCSScaler, PLS + KNN, mRMR + KNN y whole + SVC
- PLS + KNNSScaler, PCA + KNNSScaler, PLS + SVCSScaler, PLS + KNN, mRMR + KNN, whole + SVC, PLS + SVC, mRMR + KNNSScaler, PCA + SVC, mRMR + SVC, whole + SVCSScaler, mRMR + SVCSScaler, whole + LR y PLS + LR
- PCA + KNNSScaler, PLS + SVCSScaler, PLS + KNN, mRMR + KNN, whole + SVC, PLS + SVC, mRMR + KNNSScaler, PCA + SVC, mRMR + SVC, whole + SVCSScaler, mRMR + SVCSScaler, whole + LR, PLS + LR y PCA + SVCSScaler
- mRMR + KNN, whole + SVC, PLS + SVC, mRMR + KNNSScaler, PCA + SVC, mRMR + SVC, whole + SVCSScaler, mRMR + SVCSScaler, whole + LR, PLS + LR, PCA + SVCSScaler y whole + KNN
- mRMR + KNNSScaler, PCA + SVC, mRMR + SVC, whole + SVCSScaler, mRMR + SVCSScaler, whole + LR, PLS + LR, PCA + SVCSScaler, whole + KNN y whole + KNNSScaler.
- PCA + SVC, mRMR + SVC, whole + SVCSScaler, mRMR + SVCSScaler, whole + LR, PLS + LR, PCA + SVCSScaler, whole + KNN, whole + KNNSScaler, PCA + KNN y mRMR + LR
- whole + KNN, whole + KNNSScaler, PCA + KNN, mRMR + LR y PCA + LR.

Todas las demás diferencias son significativas.

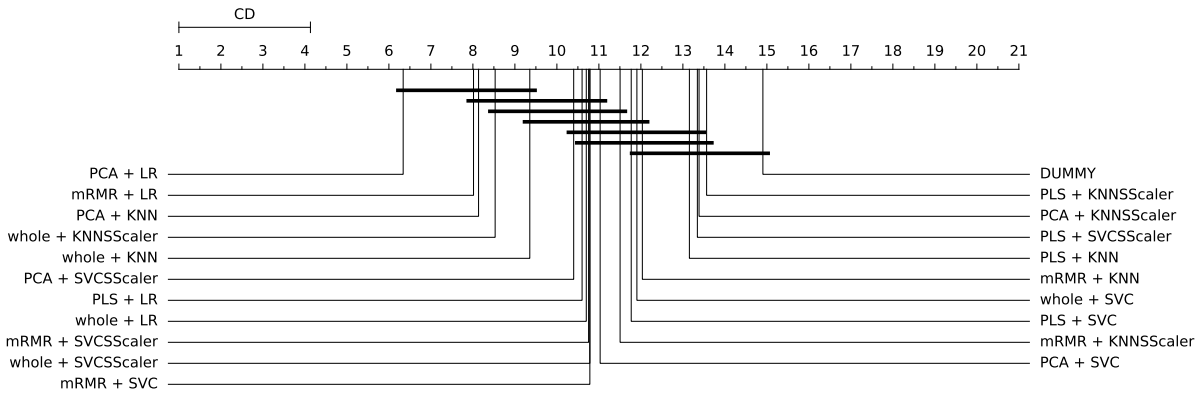


Figura 1: El diagrama CD representa los resultados del test Nemenyi para el ranking basado en el ROC AUC de los modelos con el dataset CC. Las líneas horizontales representan diferencias no significativas

2. Resultados DCOR

El análisis estadístico se hizo entre 21 modelos (poblaciones) con 100 muestras emparejadas.

El nivel de significancia para los test globales (*family-wise*) es de $\alpha = 0,050$.

No hemos podido rechazar la hipótesis nula de que la población sea normal para ninguna población (mínimo $p = 0,0526$). Por lo tanto asumimos que las poblaciones son normales.

Hemos aplicado el test Bartlett para evaluar la homogeneidad de los datos y rechazamos la hipótesis nula ($p \approx 0$) de que los datos son homocedásticos. Por lo que asumimos que los datos son heteroscedásticos.

Dado que tenemos más de dos poblaciones, donde todas son normales pero heterocedásticas, usamos el test no paramétrico de Friedman como test *omnibus* para determinar si hay alguna diferencia significativa entre las medias de las poblaciones. Utilizamos el test de Nemenyi como *post-hoc* para saber que diferencias son significativas. Mostramos, en la tabla 2, la media (M), la desviación estándar (SD), y el ranking medio (MR) de cada modelo a partir de todas las muestras. Consideraremos significativas la diferencias entre modelos si la diferencia entre rankings medios es mayor que la distancia crítica de $CD = 3,132$ dada por el test de Nemenyi.

Rechazamos la hipótesis nula ($p = 2,41 \times 10^{-37}$) del test de Friedman de que no hay diferencia entre las medidas de tendencia central de los modelos. Por lo que asumimos que hay una diferencia estadísticamente significativa entre las medianas de las poblaciones.

	MR	M	SD	CI	d	Magnitudo
PCA + SVC	15.965	0.452	0.053	[0.435, 0.469]	0.000	negligible
whole + SVCSScaler	14.190	0.471	0.051	[0.455, 0.487]	-0.363	small
whole + SVC	13.420	0.476	0.061	[0.457, 0.495]	-0.418	small
PCA + LR	13.165	0.475	0.055	[0.457, 0.492]	-0.420	small
PCA + KNN	12.720	0.482	0.048	[0.467, 0.497]	-0.586	medium
PLS + SVC	11.755	0.489	0.054	[0.472, 0.506]	-0.689	medium
PLS + KNN	11.320	0.491	0.055	[0.474, 0.508]	-0.716	medium
whole + KNN	11.015	0.494	0.052	[0.477, 0.510]	-0.794	medium
PCA + SVCSScaler	10.815	0.497	0.063	[0.477, 0.517]	-0.773	medium
PLS + SVCSScaler	10.660	0.498	0.060	[0.479, 0.517]	-0.808	large
PLS + KNNSScaler	10.650	0.497	0.053	[0.481, 0.514]	-0.850	large
whole + KNNSScaler	10.570	0.501	0.050	[0.485, 0.517]	-0.950	large
whole + LR	10.555	0.499	0.046	[0.485, 0.513]	-0.948	large
DUMMY	10.255	0.500	0.000	[0.500, 0.500]	-1.282	large
mRMR + SVC	10.225	0.502	0.051	[0.487, 0.518]	-0.970	large
mRMR + KNNSScaler	10.125	0.502	0.055	[0.485, 0.520]	-0.934	large
PLS + LR	9.965	0.503	0.067	[0.482, 0.523]	-0.837	large
mRMR + KNN	9.740	0.505	0.055	[0.488, 0.522]	-0.979	large
mRMR + SVCSScaler	9.375	0.508	0.061	[0.489, 0.527]	-0.984	large
PCA + KNNSScaler	8.485	0.518	0.058	[0.500, 0.536]	-1.186	large
mRMR + LR	6.030	0.538	0.055	[0.521, 0.555]	-1.592	large

Tabla 2: Comparativa del ROC AUC de los modelos con el dataset DCOR

Basándonos en el test *post-hoc* de Nemenyi, asumimos que no hay diferencia significativa dentro de los siguientes grupos:

- PCA + SVC, whole + SVCSScaler, whole + SVC y PCA + LR
- whole + SVCSScaler, whole + SVC, PCA + LR, PCA + KNN, PLS + SVC y PLS + KNN
- whole + SVC, PCA + LR, PCA + KNN, PLS + SVC, PLS + KNN, whole + KNN, PCA + SVCSScaler, PLS + SVCSScaler, PLS + KNNSScaler, whole + KNNSScaler y whole + LR

- PCA + LR, PCA + KNN, PLS + SVC, PLS + KNN, whole + KNN, PCA + SVCSScaler, PLS + SVCSScaler, PLS + KNNSScaler, whole + KNNSScaler, whole + LR, DUMMY, mRMR + SVC y mRMR + KNNSScaler
- PCA + KNN, PLS + SVC, PLS + KNN, whole + KNN, PCA + SVCSScaler, PLS + SVCSScaler, PLS + KNNSScaler, whole + KNNSScaler, whole + LR, DUMMY, mRMR + SVC, mRMR + KNNSScaler, PLS + LR y mRMR + KNN
- PLS + SVC, PLS + KNN, whole + KNN, PCA + SVCSScaler, PLS + SVCSScaler, PLS + KNNSScaler, whole + KNNSScaler, whole + LR, DUMMY, mRMR + SVC, mRMR + KNNSScaler, PLS + LR, mRMR + KNN y mRMR + SVCSScaler
- PLS + KNN, whole + KNN, PCA + SVCSScaler, PLS + SVCSScaler, PLS + KNNSScaler, whole + KNNSScaler, whole + LR, DUMMY, mRMR + SVC, mRMR + KNNSScaler, PLS + LR, mRMR + KNN, mRMR + SVCSScaler y PCA + KNNSScaler
- PCA + KNNSScaler and mRMR + LR.

Todas las demás diferencias son significativas.

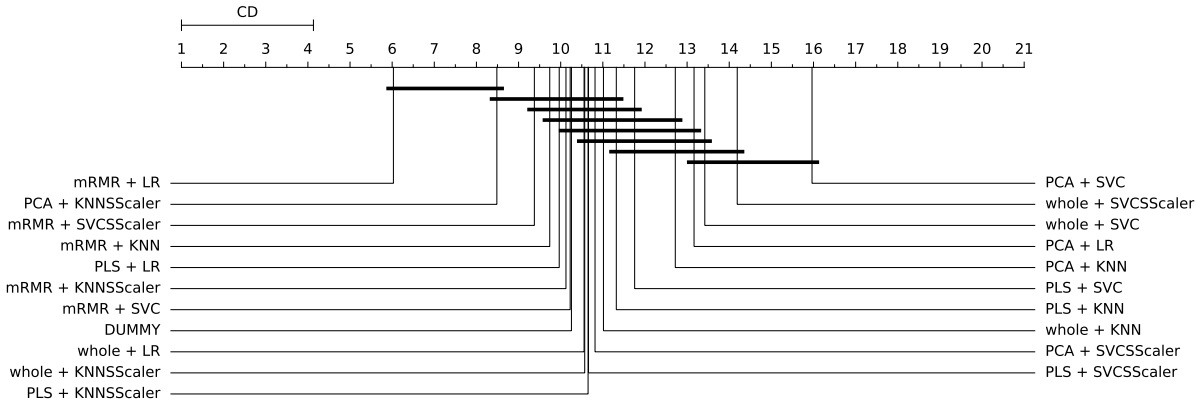


Figura 2: El diagrama CD representa los resultados del test Nemenyi para el ranking basado en el ROC AUC de los modelos con el dataset DCOR. Las líneas horizontales representan diferencias no significativas

3. Resultados FFT

El análisis estadístico se hizo entre 5 modelos (poblaciones) con 100 muestras emparejadas.

El nivel de significancia para los test globales (*family-wise*) es de $\alpha = 0,050$.

No hemos podido rechazar la hipótesis nula de que la población sea normal para ninguna población (mínimo $p = 0,013$). Por lo tanto asumimos que las poblaciones son normales.

Hemos aplicado el test Bartlett para evaluar la homogeneidad de los datos y rechazamos la hipótesis nula ($p \approx 0$) de que los datos son homocedásticos. Por lo que asumimos que los datos son heterocedásticos

Dado que tenemos más de dos poblaciones, donde todas son normales pero heterocedásticas, usamos el test no paramétrico de Friedman como test *omnibus* para determinar si hay alguna

diferencia significativa entre las medias de las poblaciones. Utilizamos el test de Nemenyi como *post-hoc* para saber que diferencias son significativas. Mostramos, en la tabla 3, la media (M), la desviación estándar (SD), y el ranking medio (MR) de cada modelo a partir de todas las muestras. Consideraremos significativas la diferencias entre modelos si la diferencia entre rankings medios es mayor que la distancia crítica de $CD = 3,807$ dada por el test de Nemenyi.

Rechazamos la hipótesis nula ($p = 4,21 \times 10^{-157}$) del test de Friedman de que no hay diferencia entre las medidas de tendencia central de los modelos. Por lo que asumimos que hay una diferencia estadísticamente significativa entre las medianas de las poblaciones.

	MR	M	SD	CI	\$d\$	Magnitude
DUMMY	21.590	0.500	0.000	[0.500, 0.500]	-	large
whole + SVC	21.590	0.500	0.000	[0.500, 0.500]	-	large
PCA + SVC	19.170	0.505	0.066	[0.484, 0.526]	-0.106	negligible
PLS + KNNSScaler	16.815	0.537	0.051	[0.521, 0.553]	-1.018	large
whole + LR	15.330	0.547	0.054	[0.530, 0.564]	-1.244	large
PCA + SVCSScaler	15.310	0.547	0.047	[0.532, 0.562]	-1.434	large
PCA + KNNSScaler	14.700	0.552	0.052	[0.536, 0.569]	-1.414	large
mRMR + KNN	14.215	0.558	0.059	[0.540, 0.577]	-1.401	large
PLS + KNN	13.820	0.560	0.052	[0.543, 0.576]	-1.623	large
PLS + SVCSScaler	13.620	0.561	0.057	[0.543, 0.579]	-1.525	large
whole + KNNSScaler	13.440	0.562	0.056	[0.544, 0.580]	-1.567	large
PCA + KNN	13.400	0.563	0.057	[0.545, 0.581]	-1.554	large
PLS + LRSScaler	12.950	0.568	0.056	[0.550, 0.586]	-1.713	large
mRMR + SVC	12.910	0.565	0.047	[0.550, 0.580]	-1.948	large
whole + LRSScaler	12.785	0.566	0.059	[0.548, 0.585]	-1.598	large
PCA + LR	12.190	0.573	0.049	[0.558, 0.589]	-2.139	large
whole + KNN	12.170	0.573	0.051	[0.557, 0.589]	-2.041	large
PCA + LRSScaler	12.085	0.573	0.051	[0.557, 0.589]	-2.026	large
mRMR + KNNSScaler	11.585	0.581	0.059	[0.563, 0.600]	-1.952	large
PLS + SVC	11.160	0.581	0.059	[0.562, 0.600]	-1.932	large
PLS + LR	11.115	0.582	0.054	[0.565, 0.599]	-2.149	large
whole + SVCSScaler	7.660	0.612	0.051	[0.595, 0.628]	-3.101	large
mRMR + LR	5.360	0.633	0.054	[0.616, 0.650]	-3.466	large
mRMR + LRSScaler	5.115	0.635	0.054	[0.618, 0.653]	-3.512	large
mRMR + SVCSScaler	4.915	0.636	0.061	[0.617, 0.655]	-3.155	large

Tabla 3: Comparativa del ROC AUC de los modelos con el dataset FFT

Basándonos en el test *post-hoc* de Nemenyi, asumimos que no hay diferencia significativa dentro de los siguientes grupos:

- DUMMY, whole + SVC y PCA + SVC
- PCA + SVC and PLS + KNNSScaler
- PLS + KNNSScaler, whole + LR, PCA + SVCSScaler, PCA + KNNSScaler, mRMR + KNN, PLS + KNN, PLS + SVCSScaler, whole + KNNSScaler y PCA + KNN
- whole + LR, PCA + SVCSScaler, PCA + KNNSScaler, mRMR + KNN, PLS + KNN, PLS + SVCSScaler, whole + KNNSScaler, PCA + KNN, PLS + LRSScaler, mRMR + SVC, whole + LRSScaler, PCA + LR, whole + KNN, PCA + LRSScaler y mRMR + KNNSScaler

- PCA + KNNSScaler, mRMR + KNN, PLS + KNN, PLS + SVCSScaler, whole + KNNSScaler, PCA + KNN, PLS + LRSScaler, mRMR + SVC, whole + LRSScaler, PCA + LR, whole + KNN, PCA + LRSScaler, mRMR + KNNSScaler, PLS + SVC y PLS + LR
- PLS + SVC, PLS + LR y whole + SVCSScaler
- whole + SVCSScaler, mRMR + LR, mRMR + LRSScaler y mRMR + SVCSScaler.

Todas las demás diferencias son significativas.

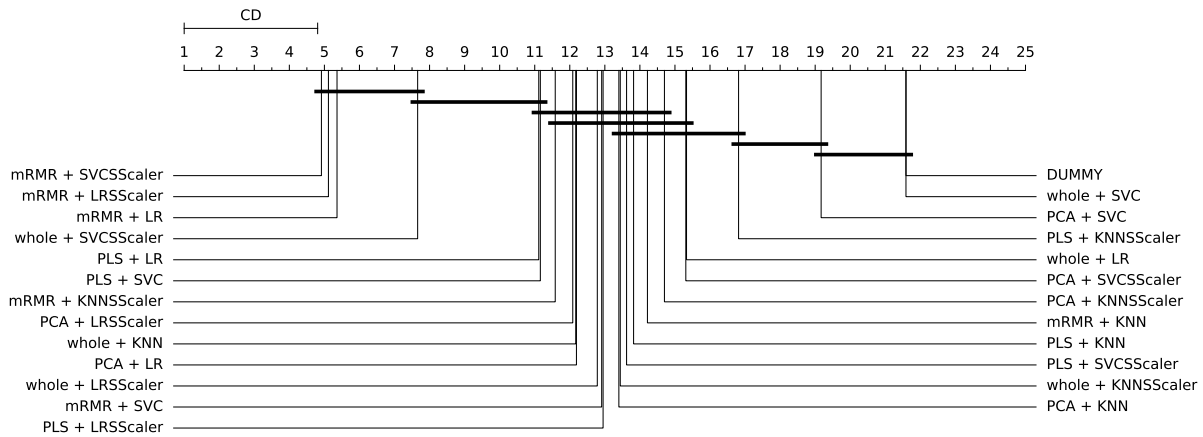


Figura 3: El diagrama CD representa los resultados del test Nemenyi para el ranking basado en el ROC AUC de ls modelos con el dataset FFT. Las líneas horizontales representan diferencias no significativas