

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 $p = 0,017$). 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	17.485	0.500	0.000	[0.500, 0.500]	-	large
PCA + KNNSScaler	14.575	0.523	0.072	[0.501, 0.546]	-0.459	small
PLS + KNNSScaler	14.180	0.530	0.072	[0.507, 0.552]	-0.583	medium
whole + SVC	12.915	0.553	0.065	[0.532, 0.573]	-1.149	large
PCA + SVC	12.520	0.552	0.074	[0.529, 0.575]	-1.007	large
PCA + SVCSScaler	12.210	0.553	0.091	[0.525, 0.582]	-0.831	large
PLS + KNN	11.870	0.560	0.075	[0.537, 0.583]	-1.134	large
PLS + SVCSScaler	11.640	0.561	0.072	[0.539, 0.583]	-1.199	large
whole + KNN	11.185	0.566	0.076	[0.542, 0.590]	-1.234	large
whole + KNNSScaler	10.880	0.572	0.073	[0.549, 0.594]	-1.394	large
PCA + KNN	10.860	0.569	0.070	[0.547, 0.591]	-1.384	large
PLS + LR	10.640	0.570	0.070	[0.548, 0.592]	-1.411	large
PLS + SVC	10.430	0.575	0.076	[0.551, 0.599]	-1.392	large
PCA + LR	10.105	0.578	0.069	[0.557, 0.600]	-1.616	large
whole + SVCSScaler	9.845	0.581	0.070	[0.559, 0.603]	-1.640	large
mRMR + KNN	8.615	0.596	0.081	[0.571, 0.621]	-1.673	large
whole + LR	8.590	0.591	0.075	[0.567, 0.614]	-1.713	large
mRMR + SVCSScaler	8.495	0.600	0.083	[0.574, 0.626]	-1.703	large
mRMR + SVC	8.205	0.603	0.082	[0.577, 0.629]	-1.772	large
mRMR + LR	7.940	0.605	0.067	[0.585, 0.626]	-2.229	large
mRMR + KNNSScaler	7.815	0.605	0.069	[0.583, 0.626]	-2.154	large

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

Rechazamos la hipótesis nula ($p = 6,21 \times 10^{-55}$) 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 y PCA + KNNSScaler
- PCA + KNNSScaler, PLS + KNNSScaler, whole + SVC, PCA + SVC, PCA + SVCSScaler, PLS + KNN y PLS + SVCSScaler
- PLS + KNNSScaler, whole + SVC, PCA + SVC, PCA + SVCSScaler, PLS + KNN, PLS + SVCSScaler y whole + KNN
- whole + SVC, PCA + SVC, PCA + SVCSScaler, PLS + KNN, PLS + SVCSScaler, whole + KNN, whole + KNNSScaler, PCA + KNN, PLS + LR, PLS + SVC, PCA + LR y whole + SVCSScaler
- PLS + SVCSScaler, whole + KNN, whole + KNNSScaler, PCA + KNN, PLS + LR, PLS + SVC, PCA + LR, whole + SVCSScaler, mRMR + KNN y whole + LR
- whole + KNN, whole + KNNSScaler, PCA + KNN, PLS + LR, PLS + SVC, PCA + LR, whole + SVCSScaler, mRMR + KNN, whole + LR, mRMR + SVCSScaler y mRMR + SVC
- whole + KNNSScaler, PCA + KNN, PLS + LR, PLS + SVC, PCA + LR, whole + SVCSScaler, mRMR + KNN, whole + LR, mRMR + SVCSScaler, mRMR + SVC, mRMR + LR y mRMR + KNNSScaler.

Todas las demás diferencias son significativas.

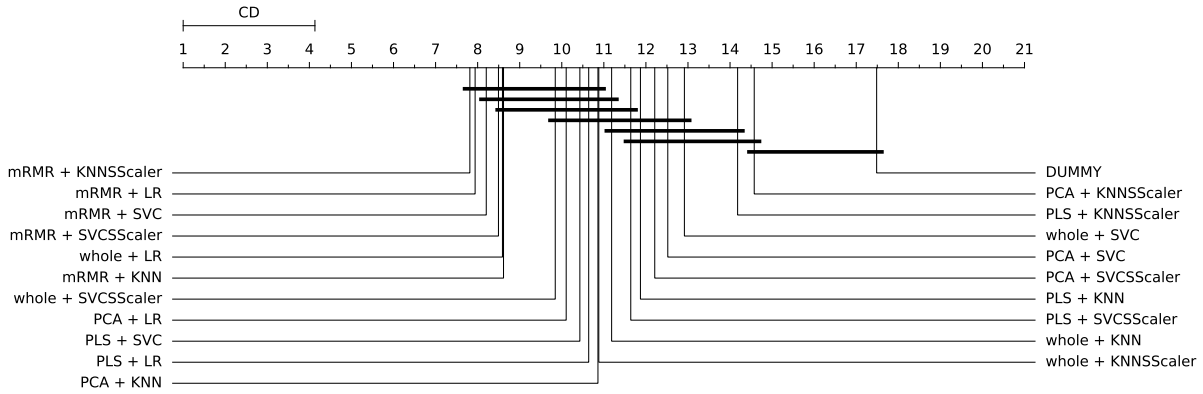


Figura 1: El diagrama CD representa los resultados del test Nemenyi para el ranking basado en el ROC AUC de los modelos con el dataset CC. Las líneas horizontales representan diferencias no significativas

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 $p = 0,031$). 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 ($p = 9,34 \times 10^{-42}$) 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
PCA + KNN	15.130	0.443	0.074	[0.419, 0.466]	0.000	negligible
whole + KNN	13.815	0.459	0.080	[0.434, 0.484]	-0.215	small
whole + KNNSScaler	13.595	0.461	0.079	[0.436, 0.485]	-0.238	small
mRMR + KNN	13.010	0.468	0.066	[0.448, 0.489]	-0.364	small
PCA + SVC	12.525	0.475	0.077	[0.451, 0.499]	-0.432	small
mRMR + KNNSScaler	12.500	0.473	0.072	[0.450, 0.495]	-0.412	small
mRMR + SVCSScaler	12.380	0.476	0.078	[0.452, 0.500]	-0.436	small
whole + SVCSScaler	11.885	0.482	0.075	[0.459, 0.506]	-0.528	medium
PLS + SVC	11.340	0.489	0.075	[0.466, 0.513]	-0.627	medium
mRMR + SVC	11.245	0.494	0.083	[0.468, 0.520]	-0.654	medium
PCA + KNNSScaler	11.210	0.495	0.090	[0.467, 0.523]	-0.633	medium
PLS + KNNSScaler	10.955	0.497	0.072	[0.474, 0.519]	-0.735	medium
PCA + SVCSScaler	10.760	0.496	0.087	[0.468, 0.523]	-0.653	medium
PLS + KNN	10.365	0.499	0.061	[0.480, 0.518]	-0.825	large
DUMMY	10.265	0.500	0.000	[0.500, 0.500]	-1.090	large
whole + SVC	9.920	0.507	0.063	[0.487, 0.526]	-0.935	large
PLS + SVCSScaler	9.620	0.507	0.077	[0.483, 0.531]	-0.849	large
whole + LR	8.995	0.516	0.060	[0.497, 0.535]	-1.084	large
PCA + LR	7.895	0.535	0.083	[0.509, 0.561]	-1.176	large
PLS + LR	6.845	0.547	0.086	[0.520, 0.573]	-1.298	large
mRMR + LR	6.745	0.545	0.070	[0.523, 0.567]	-1.419	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 + KNN, whole + KNN, whole + KNNSScaler, mRMR + KNN, PCA + SVC, mRMR + KNNSScaler y mRMR + SVCSScaler
- whole + KNN, whole + KNNSScaler, mRMR + KNN, PCA + SVC, mRMR + KNNSScaler, mRMR + SVCSScaler, whole + SVCSScaler, PLS + SVC, mRMR + SVC, PCA + KNNSScaler, PLS + KNNSScaler y PCA + SVCSScaler
- mRMR + KNN, PCA + SVC, mRMR + KNNSScaler, mRMR + SVCSScaler, whole +

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

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

Todas las demás diferencias son significativas.

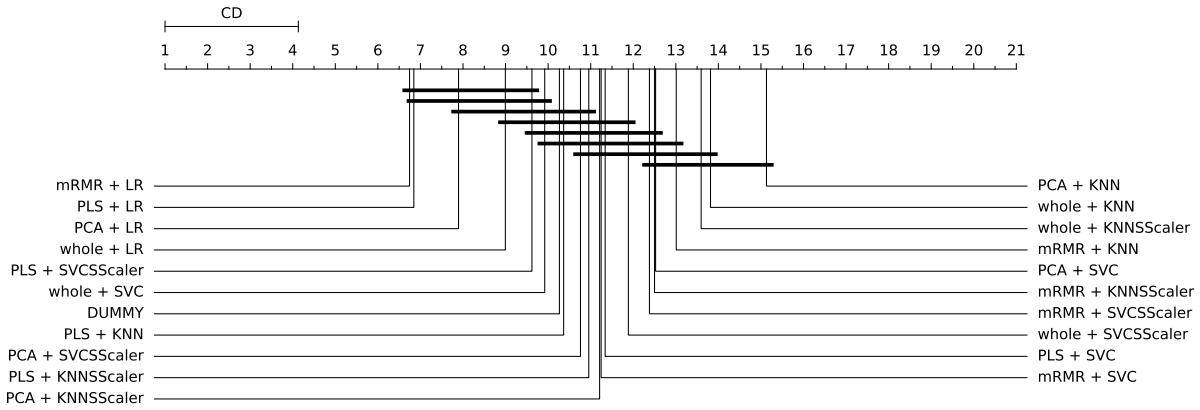


Figura 2: El diagrama CD representa los resultados del test Nemenyi para el ranking basado en el ROC AUC de los modelos con el dataset DCOR. Las líneas horizontales representan diferencias no significativas

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 $p = 0,006$). 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 ($p = 1,3 \times 10^{-128}$) 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
PCA + KNNSScaler	19.355	0.471	0.071	[0.449, 0.494]	0.000	negligible
DUMMY	18.030	0.500	0.000	[0.500, 0.500]	-0.570	medium
whole + SVC	18.030	0.500	0.000	[0.500, 0.500]	-0.570	medium
PCA + KNN	17.940	0.487	0.069	[0.465, 0.509]	-0.224	small
PCA + SVCSScaler	17.535	0.490	0.075	[0.466, 0.514]	-0.260	small
PCA + SVC	15.850	0.509	0.056	[0.492, 0.527]	-0.596	medium
PCA + LRSScaler	15.705	0.515	0.069	[0.493, 0.537]	-0.622	medium
PCA + LR	14.995	0.520	0.073	[0.497, 0.543]	-0.680	medium
mRMR + SVC	14.850	0.527	0.056	[0.510, 0.545]	-0.878	large
PLS + KNNSScaler	14.705	0.529	0.065	[0.509, 0.550]	-0.853	large
whole + KNN	14.690	0.524	0.066	[0.503, 0.545]	-0.774	medium
whole + KNNSScaler	13.700	0.533	0.068	[0.511, 0.554]	-0.886	large
PLS + KNN	12.790	0.543	0.067	[0.521, 0.564]	-1.033	large
PLS + LRSScaler	12.485	0.547	0.071	[0.525, 0.570]	-1.072	large
mRMR + KNNSScaler	12.250	0.551	0.071	[0.528, 0.574]	-1.116	large
PLS + SVCSScaler	12.240	0.551	0.075	[0.528, 0.575]	-1.097	large
PLS + SVC	12.050	0.549	0.077	[0.525, 0.574]	-1.055	large
PLS + LR	11.890	0.551	0.070	[0.529, 0.573]	-1.137	large
mRMR + KNN	10.890	0.564	0.081	[0.538, 0.589]	-1.213	large
whole + SVCSScaler	10.100	0.570	0.066	[0.549, 0.591]	-1.445	large
whole + LR	9.570	0.572	0.070	[0.550, 0.595]	-1.428	large
mRMR + SVCSScaler	6.960	0.615	0.091	[0.586, 0.644]	-1.765	large
mRMR + LR	6.385	0.621	0.085	[0.594, 0.648]	-1.917	large
whole + LRSScaler	6.340	0.608	0.077	[0.584, 0.632]	-1.853	large
mRMR + LRSScaler	5.665	0.627	0.089	[0.599, 0.656]	-1.935	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:

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

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

Todas las demás diferencias son significativas.

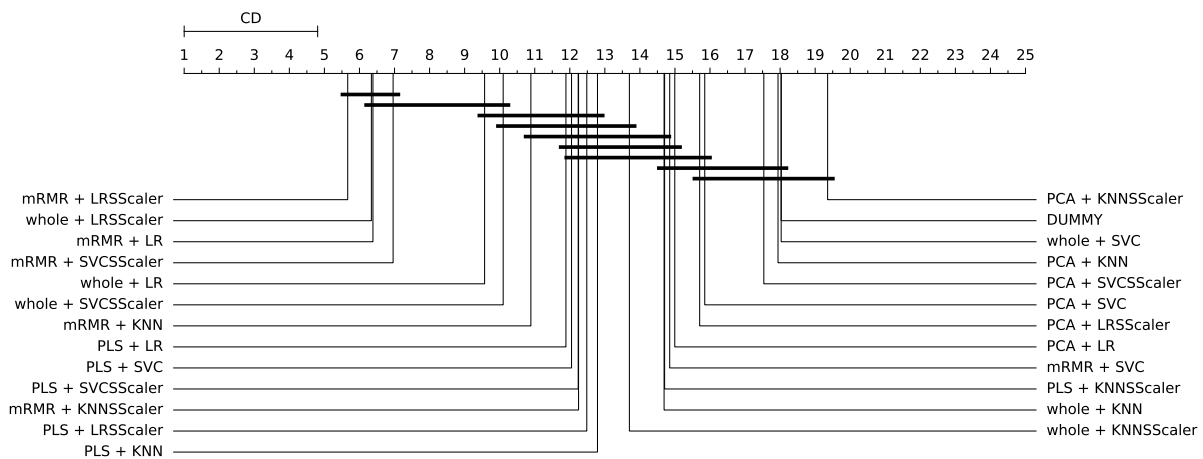


Figura 3: El diagrama CD representa los resultados del test Nemenyi para el ranking basado en el ROC AUC de los modelos con el dataset FFT. Las líneas horizontales representan diferencias no significativas