

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 $\alpha = 0,050$.

Rechazamos la hipótesis nula de que la población es normal para el modelo PLS + SVC ($p=0.002$). Por lo tanto asumimos no que todas las poblaciones son normales.

Dado que tenemos más de dos poblaciones y que al menos una de ella no es normal, usamos el test no paramétrico de Friedman como test *omnibus* para determinar si hay alguna diferencia significativa entre las medianas de las poblaciones. Utilizamos el test de Nemenyi como *post-hoc* para saber que diferencias son significativas. Mostramos, en la tabla 1, la mediana (MD), la desviación absoluta sobre a la mediana (MAD), 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	MED	MAD	CI	γ	Magnitude
DUMMY	17.340	0.500	0.000	[0.500, 0.500]	-	large
PCA + KNNSScaler	14.745	0.510	0.035	[0.493, 0.548]	-0.289	small
PLS + KNNSScaler	14.060	0.520	0.036	[0.490, 0.556]	-0.542	medium
mRMR + KNN	12.880	0.532	0.031	[0.510, 0.560]	-0.988	large
whole + SVC	12.755	0.533	0.033	[0.508, 0.562]	-0.939	large
PLS + KNN	12.640	0.522	0.036	[0.504, 0.562]	-0.583	medium
PCA + SVC	12.425	0.535	0.041	[0.506, 0.571]	-0.815	large
mRMR + KNNSScaler	12.350	0.535	0.029	[0.514, 0.563]	-1.124	large
PLS + SVC	11.545	0.548	0.028	[0.517, 0.568]	-1.654	large
PLS + SVCSScaler	11.390	0.543	0.035	[0.516, 0.573]	-1.159	large
mRMR + SVCSScaler	11.275	0.548	0.032	[0.519, 0.576]	-1.435	large
PCA + SVCSScaler	10.650	0.547	0.041	[0.520, 0.587]	-1.109	large
whole + SVCSScaler	10.475	0.547	0.031	[0.527, 0.578]	-1.445	large
whole + KNN	10.375	0.550	0.036	[0.518, 0.582]	-1.305	large
whole + KNNSScaler	10.040	0.554	0.036	[0.523, 0.587]	-1.405	large
mRMR + SVC	9.945	0.552	0.033	[0.529, 0.585]	-1.499	large
PCA + KNN	9.245	0.561	0.032	[0.533, 0.585]	-1.849	large
PLS + LR	7.950	0.569	0.033	[0.545, 0.602]	-1.960	large
mRMR + LR	7.520	0.576	0.033	[0.547, 0.601]	-2.180	large
whole + LR	7.070	0.578	0.030	[0.555, 0.598]	-2.456	large
PCA + LR	4.325	0.604	0.032	[0.570, 0.623]	-3.115	large

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

Rechazamos la hipótesis nula ($p = 3,86 \times 10^{-78}$) 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, mRMR + KNN, whole + SVC, PLS + KNN, PCA + SVC y mRMR + KNNSScaler
- PLS + KNNSScaler, mRMR + KNN, whole + SVC, PLS + KNN, PCA + SVC, mRMR + KNNSScaler, PLS + SVC, PLS + SVCSScaler y mRMR + SVCSScaler
- mRMR + KNN, whole + SVC, PLS + KNN, PCA + SVC, mRMR + KNNSScaler, PLS + SVC, PLS + SVCSScaler, mRMR + SVCSScaler, PCA + SVCSScaler, whole + SVCSScaler, whole + KNN, whole + KNNSScaler y mRMR + SVC
- mRMR + KNNSScaler, PLS + SVC, PLS + SVCSScaler, mRMR + SVCSScaler, PCA + SVCSScaler, whole + SVCSScaler, whole + KNN, whole + KNNSScaler, mRMR + SVC y PCA + KNN
- PCA + SVCSScaler, whole + SVCSScaler, whole + KNN, whole + KNNSScaler, mRMR + SVC, PCA + KNN, PLS + LR y mRMR + LR
- whole + KNNSScaler, mRMR + SVC, PCA + KNN, PLS + LR, mRMR + LR y whole + LR
- whole + LR y PCA + LR

Todas las demás diferencias son significativas.

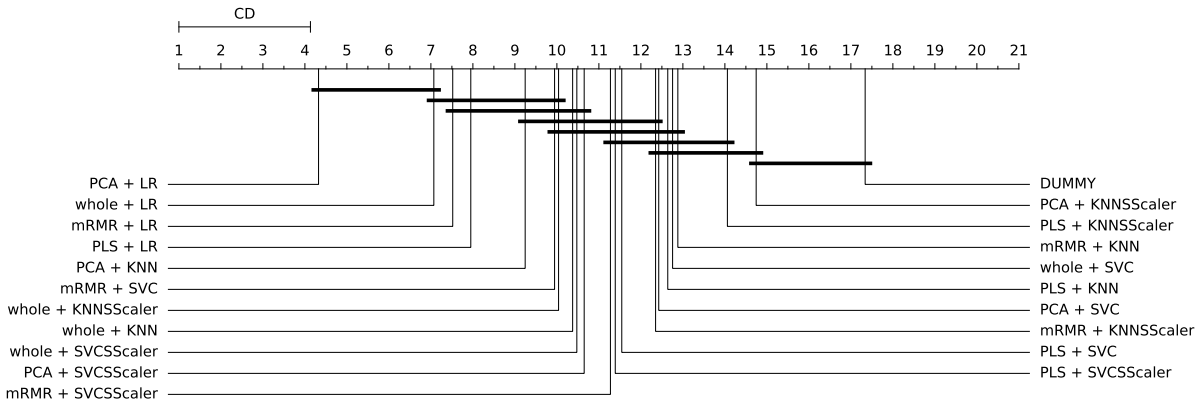


Figura 1: El diagrama CD representa los resultados del test Nemenyi para el ranking basado en el ROC AUC de los modelos con el dataset CC. Las líneas horizontales representan diferencias no significativas

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 $p = 0,083$). 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 ($p = 3,12 \times 10^{-27}$) 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 + SVC	14.050	0.465	0.045	[0.451, 0.479]	0.000	negligible
whole + SVCSScaler	13.775	0.468	0.060	[0.450, 0.487]	-0.061	negligible
PCA + SVCSScaler	13.085	0.474	0.044	[0.460, 0.488]	-0.193	negligible
whole + KNN	13.060	0.472	0.054	[0.455, 0.489]	-0.139	negligible
whole + KNNSScaler	12.970	0.475	0.047	[0.460, 0.490]	-0.207	small
whole + SVC	12.870	0.475	0.049	[0.460, 0.490]	-0.207	small
PCA + KNN	12.145	0.479	0.050	[0.463, 0.494]	-0.287	small
mRMR + SVCSScaler	11.520	0.483	0.053	[0.467, 0.500]	-0.370	small
mRMR + KNNSScaler	11.485	0.486	0.050	[0.470, 0.502]	-0.437	small
mRMR + KNN	11.285	0.487	0.056	[0.469, 0.504]	-0.424	small
PCA + LR	11.090	0.488	0.040	[0.475, 0.500]	-0.526	medium
mRMR + SVC	11.085	0.488	0.051	[0.472, 0.504]	-0.477	small
PLS + SVC	9.940	0.495	0.055	[0.478, 0.512]	-0.596	medium
PLS + KNN	9.820	0.499	0.052	[0.482, 0.515]	-0.689	medium
PLS + KNNSScaler	9.675	0.500	0.052	[0.484, 0.516]	-0.720	medium
DUMMY	9.525	0.500	0.000	[0.500, 0.500]	-1.099	large
whole + LR	9.385	0.501	0.044	[0.487, 0.515]	-0.804	large
PCA + KNNSScaler	9.145	0.502	0.050	[0.486, 0.517]	-0.762	medium
PLS + SVCSScaler	8.465	0.511	0.051	[0.495, 0.527]	-0.949	large
mRMR + LR	8.395	0.509	0.046	[0.494, 0.523]	-0.960	large
PLS + LR	8.230	0.512	0.051	[0.496, 0.528]	-0.982	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, PCA + SVCSScaler, whole + KNN, whole + KNNSScaler, whole + SVC, PCA + KNN, mRMR + SVCSScaler, mRMR + KNNSScaler, mRMR + KNN, PCA + LR y mRMR + SVC
- whole + KNN, whole + KNNSScaler, whole + SVC, PCA + KNN, mRMR + SVCSScaler,

mRMR + KNNSScaler, mRMR + KNN, PCA + LR, mRMR + SVC y PLS + SVC

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

Todas las demás diferencias son significativas.

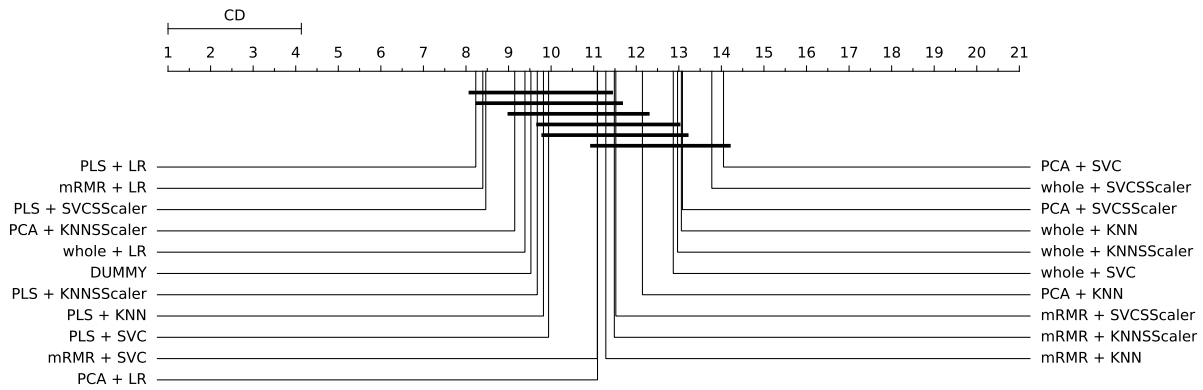


Figura 2: El diagrama CD representa los resultados del test Nemenyi para el ranking basado en el ROC AUC de los modelos con el dataset DCOR. Las líneas horizontales representan diferencias no significativas

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 $p = 0,0028$). 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 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 ($p = 5,91 \times 10^{-102}$) 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	19.680	0.500	0.000	[0.500, 0.500]	-	large
whole + SVC	19.680	0.500	0.000	[0.500, 0.500]	-	large
PCA + SVC	18.090	0.504	0.049	[0.489, 0.520]	-0.130	negligible
whole + LR	17.770	0.506	0.048	[0.491, 0.521]	-0.175	negligible
PLS + KNNSScaler	16.050	0.520	0.049	[0.504, 0.535]	-0.568	medium
mRMR + KNN	15.315	0.527	0.049	[0.511, 0.543]	-0.774	medium
whole + KNN	15.000	0.528	0.047	[0.513, 0.543]	-0.849	large
PCA + KNN	14.270	0.533	0.048	[0.518, 0.549]	-0.973	large
mRMR + KNNSScaler	14.100	0.535	0.050	[0.519, 0.551]	-0.993	large
PCA + KNNSScaler	14.060	0.534	0.047	[0.519, 0.549]	-1.037	large
PLS + KNN	13.215	0.544	0.054	[0.527, 0.561]	-1.145	large
PCA + LRSScaler	13.085	0.545	0.057	[0.527, 0.563]	-1.115	large
PLS + SVCSScaler	13.080	0.542	0.054	[0.525, 0.559]	-1.121	large
mRMR + SVC	12.640	0.545	0.051	[0.529, 0.561]	-1.242	large
PCA + LR	11.915	0.553	0.055	[0.536, 0.571]	-1.375	large
whole + KNNSScaler	11.885	0.546	0.055	[0.529, 0.564]	-1.191	large
PLS + SVC	11.390	0.554	0.058	[0.536, 0.573]	-1.328	large
PLS + LRSScaler	10.840	0.559	0.054	[0.541, 0.576]	-1.524	large
whole + LRSScaler	10.785	0.553	0.058	[0.535, 0.571]	-1.296	large
PLS + LR	10.785	0.556	0.053	[0.539, 0.573]	-1.503	large
PCA + SVCSScaler	10.615	0.555	0.053	[0.539, 0.572]	-1.481	large
mRMR + SVCSScaler	8.530	0.574	0.050	[0.558, 0.590]	-2.075	large
mRMR + LRSScaler	7.950	0.579	0.049	[0.563, 0.595]	-2.252	large
mRMR + LR	7.400	0.583	0.046	[0.568, 0.598]	-2.539	large
whole + SVCSScaler	6.870	0.585	0.047	[0.570, 0.600]	-2.577	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, PCA + SVC, whole + LR y PLS + KNNSScaler
- PCA + SVC, whole + LR, PLS + KNNSScaler, mRMR + KNN y whole + KNN
- whole + LR, PLS + KNNSScaler, mRMR + KNN, whole + KNN, PCA + KNN, mRMR + KNNSScaler y PCA + KNNSScaler
- PLS + KNNSScaler, mRMR + KNN, whole + KNN, PCA + KNN, mRMR + KNNSScaler, PCA + KNNSScaler, PLS + KNN, PCA + LRSScaler, PLS + SVCSScaler y mRMR + SVC

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

Todas las demás diferencias son significativas.

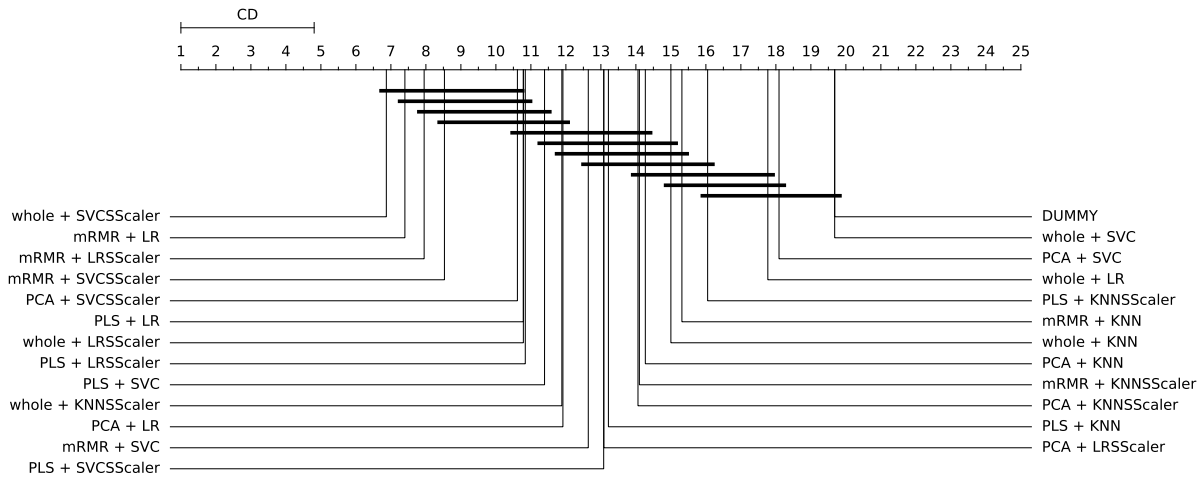


Figura 3: El diagrama CD representa los resultados del test Nemenyi para el ranking basado en el ROC AUC de los modelos con el dataset FFT. Las líneas horizontales representan diferencias no significativas