REPORT-4

cfunHDDC with Fourier Basis

Initializer	Threshold	Alphamin	Eta1	Eta2	CCR	Class 1 Outliers	Class 2 Outliers
kmeans	0.001	0.5	4.851801	12.54592	0.76	27/67	106/133
		0.6	4.851801	12.54592	0.76	27/67	106/133
		0.7	4.881624	12.56158	0.76	27/67	106/133
		0.8	5.192161	12.9742	0.76	28/67	110/133
		0.85	5.634242	12.22906	0.76	29/67	113/133
		0.9	5.775491	12.1989	0.76	33/67	113/133
		0.95	7.390356	12.47509	0.755	39/67	113/133
	0.01	0.5	4.916504	12.98552	0.745	26/67	107/133
		0.6	4.916504	12.98552	0.745	26/67	107/133
		0.7	4.941505	13.02134	0.745	26/67	107/133
		0.8	5.099383	13.04572	0.745	27/67	107/133
		0.85	5.117947	12.64636	0.745	29/67	110/133
		0.9	6.287614	11.33169	0.75	31/67	110/133
		0.95	5.20952	13.98913	0.765	48/67	130/133
	0.05	0.5	3.660024	15.92869	0.74	19/67	108/133
		0.6	3.714448	15.7564	0.74	19/67	108/133
		0.7	3.807025	15.36336	0.74	21/67	110/133
		0.8	4.336278	14.6707	0.745	28/67	113/133
		0.85	4.698932	14.18823	0.735	23/67	110/133
		0.9	4.896812	13.49985	0.735	27/67	111/133
		0.95	12.96076	14.18095	0.75	38/67	117/133

	0.1	0.5	13.83218	4.715653	0.735	20/67	108/133
		0.6	13.77458	4.755339	0.735	20/67	108/133
		0.7	13.73522	4.932894	0.725	24/67	110/133
		0.8	9.991508	14.6662	0.745	36/67	117/133
		0.85	9.959099	14.62584	0.745	36/67	117/133
		0.9	9.880587	14.43394	0.745	38/67	117/133
		0.95	14.60717	3.738225	0.735	36/67	117/133
	0.2	0.5	12.09131	5.371083	0.755	25/67	108/133
		0.6	12.10309	5.376286	0.755	25/67	108/133
		0.7	12.14522	5.319775	0.755	25/67	108/133
		0.8	12.09836	4.748381	0.755	30/67	113/133
		0.85	4.780299	13.42369	0.75	27/67	111/133
		0.9	4.10241	19.68735	0.76	46/67	129/133
		0.95	4.615807	19.27739	0.765	50/67	132/133
	0.3	0.5	5.417174	21.87661	0.74	19/67	101/133
		0.6	4.373444	24.07831	0.735	23/67	104/133
		0.7	4.403296	23.9974	0.735	23/67	104/133
		0.8	4.392968	23.83858	0.74	23/67	105/133
		0.85	3.829425	23.60799	0.74	27/67	110/133
		0.9	4.80912	22.14721	0.75	29/67	113/133
		0.95	5.93394	3.494555	0.55	48/67	127/133
	0.4	0.5	5.417174	21.87661	0.74	19/67	101/133
		0.6	4.373444	24.07831	0.735	23/67	104/133
		0.7	4.403296	23.9974	0.735	23/67	104/133
		0.8	4.392968	23.83858	0.74	23/67	105/133

		0.85	3.829425	23.60799	0.74	27/67	110/133
		0.9	4.80912	22.14721	0.75	29/67	113/133
		0.95	9.864785	5.109795	0.75	36/67	123/133
random	0.001	0.5	7.340147	8.721371	0.77	26/67	105/133
		0.6	7.340147	8.721371	0.77	26/67	105/133
		0.7	7.360263	8.758391	0.77	26/67	105/133
		0.8	8.944482	7.134026	0.78	50/67	123/133
		0.85	5.628841	6.529131	0.79	47/67	120/133
		0.9	7.463936	4.749399	0.795	54/67	122/133
		0.95	4.902774	7.874592	0.795	54/67	125/133
	0.01	0.5	8.915587	6.530792	0.77	26/67	105/133
		0.6	5.981619	9.654887	0.765	23/67	105/133
		0.7	7.085553	7.762559	0.84	48/67	111/133
		0.8	7.076365	7.861789	0.84	48/67	113/133
		0.85	7.059114	7.93761	0.84	48/67	113/133
		0.9	7.092878	8.034011	0.835	48/67	114/133
		0.95	6.553035	6.591264	0.795	58/67	125/133
	0.05	0.5	9.821098	6.041657	0.775	23/67	104/133
		0.6	9.820287	6.045162	0.775	23/67	104/133
		0.7	9.803387	6.197766	0.77	23/67	106/133
		0.8	10.99151	4.193703	0.79	29/67	103/133
		0.85	10.94827	4.229101	0.79	30/67	103/133
		0.9	72.10354	2.316662	0.765	51/67	132/133
		0.95	3.489963	31.59507	0.745	54/67	131/133
	0.1	0.5	2.941184	151.8979	0.77	27/67	115/133

		0.6	2.755144	151.9469	0.77	29/67	118/133
		0.7	2.641318	152.7392	0.77	37/67	123/133
		0.8	2.605505	154.0694	0.775	43/67	128/133
		0.85	2.569878	154.3025	0.775	47/67	129/133
		0.9	2.688531	155.0571	0.775	54/67	132/133
		0.95	10.40805	49.47463	0.805	60/67	133/133
	0.2	0.5	15.09518	6.592133	0.72	12/67	96/133
		0.6	14.96766	6.472133	0.72	12/67	96/133
		0.7	3.668319	13.75549	0.745	23/67	107/133
		0.8	13.87498	3.839593	0.71	27/67	106/133
		0.85	2.994477	13.73651	0.77	40/67	120/133
		0.9	14.22036	3.03651	0.765	43/67	122/133
		0.95	8.767399	7.056145	0.785	51/67	126/133
	0.3	0.5	13.9963	3.80061	0.745	21/67	105/133
		0.6	13.86783	3.733776	0.745	22/67	107/133
		0.7	13.75549	3.668306	0.745	23/67	107/133
		0.8	13.00803	3.010589	0.785	35/67	116/133
		0.85	13.73653	2.994478	0.77	40/67	120/133
		0.9	14.22041	3.036501	0.765	43/67	122/133
		0.95	7.056141	8.767435	0.785	51/67	126/133
	0.4	0.5	7.261258	6.068487	0.785	24/67	90/133
		0.6	7.275038	6.075912	0.785	24/67	91/133
		0.7	7.31652	6.065175	0.785	24/67	92/133
		0.8	7.339745	5.971047	0.785	24/67	98/133
		0.85	6.244864	6.961841	0.825	54/67	109/133

		0.9	6.280434	6.949104	0.815	54/67	111/133
		0.95	6.304806	7.143745	0.8	54/67	116/133

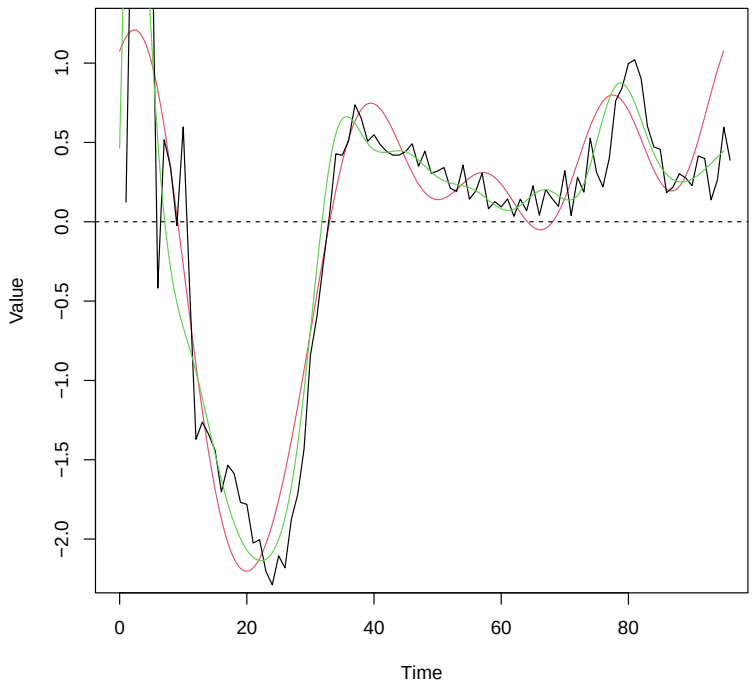
The cfunHDDC algorithm was run on the ECG data converted into a functional data object using a Fourier basis. The parameters were varied according to the table and the CCR for each configuration was recorded. The best model was chosen automatically using the BIC metric. The highest CCR was obtained with the "**Threshold**" set to **0.01**, the "**Initializer**" set to "**random**", the "**Alphamin**" set to **0.7**, **0.8** or **0.85**. This configuration was able to achieve a **CCR** of **0.84** on the data. The total number of misclassified labels are **32**. This data has a total of **22** outliers from both classes. The number of outliers that are misclassified is **4**.

Function Approximation: Fourier vs Bspline (old values)

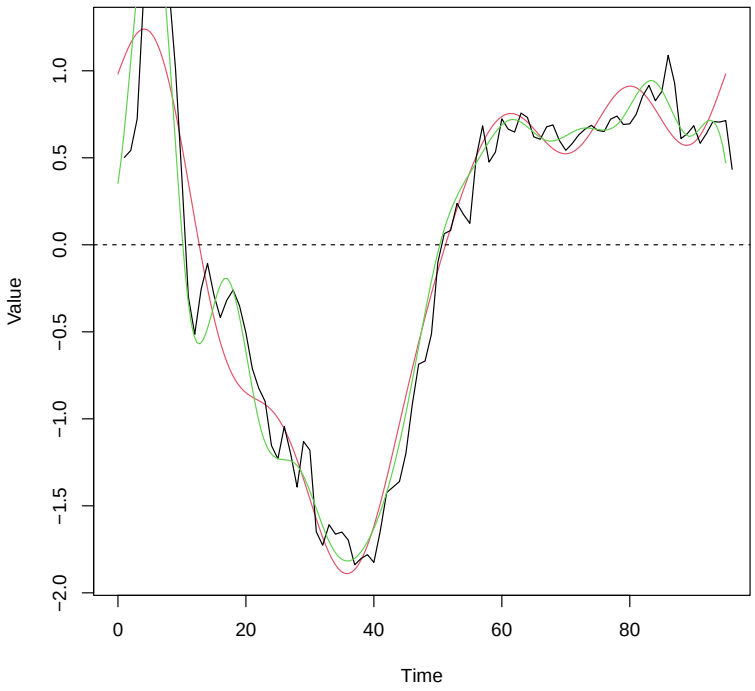
Original data => **BLACK**

Fourier basis => **RED**

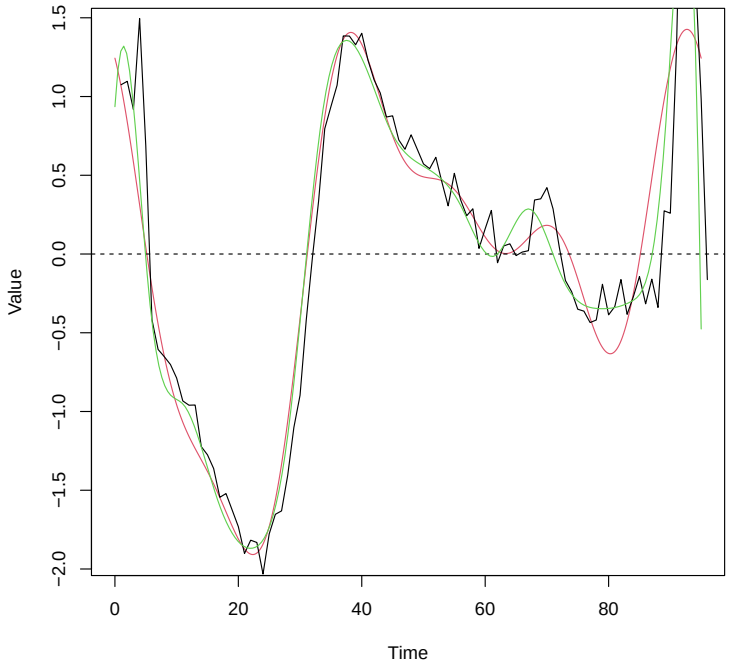
Bspline basis => **GREEN**



Feature Vector 1



Feature Vector 50



Feature Vector 150

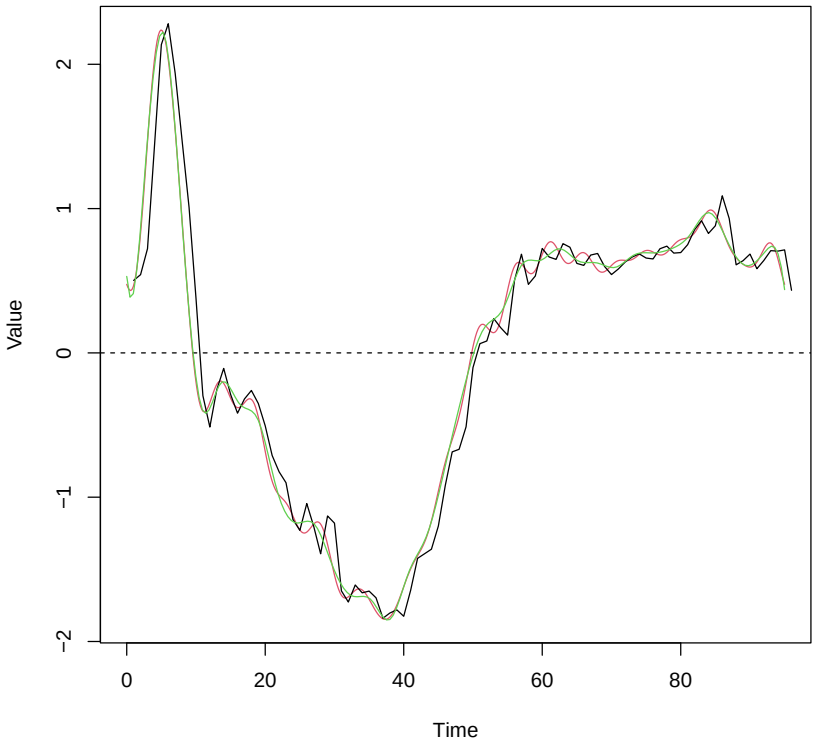
From the example plots we can see that the Bspline basis (**20 splines**) does a better job of approximation of the original data than the Fourier basis. However, it is interesting to note that even though Fourier basis (**11 splines**) does not approximate well, it consistently provides the better results while clustering.

Function Approximation: Fourier vs Bspline (new values)

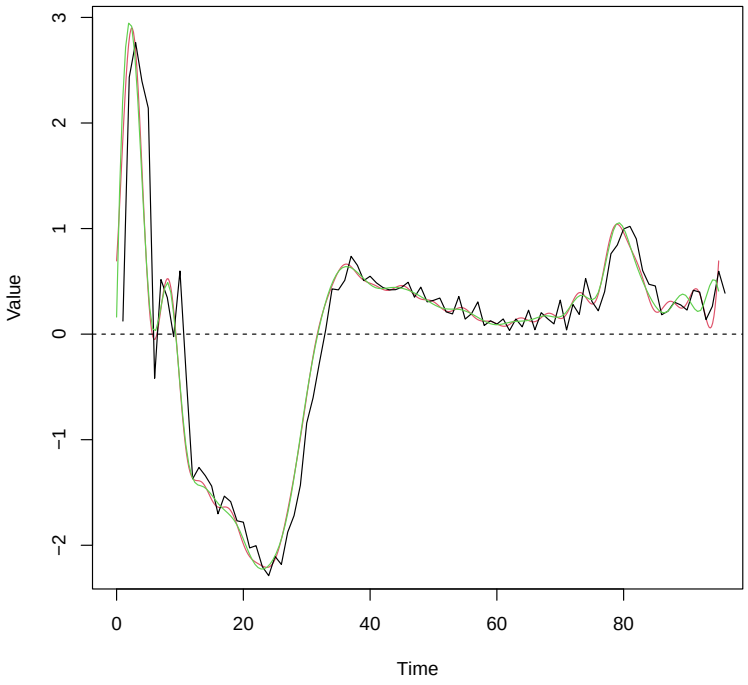
Original data => **BLACK**

Fourier basis => **RED**

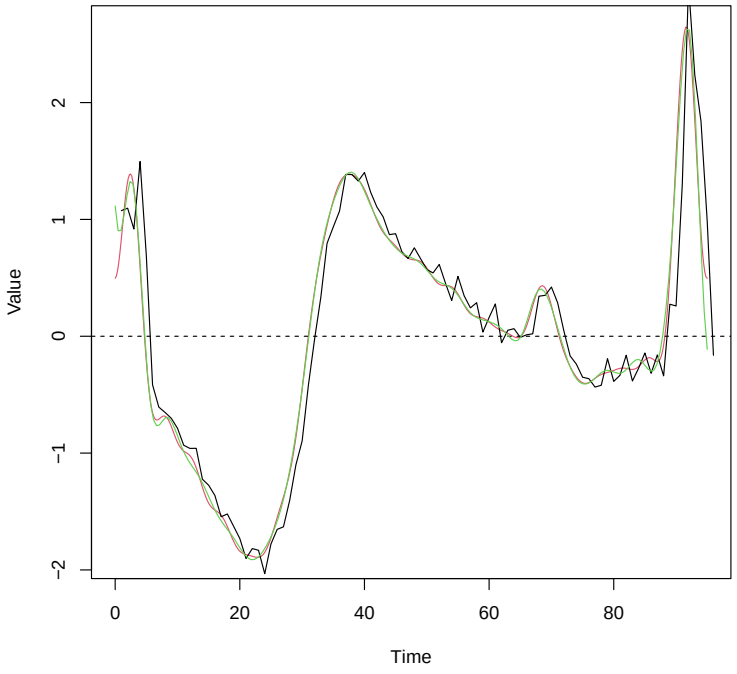
Bspline basis => **GREEN**



Feature Vector 1



Feature Vector 50



Feature Vector 150

From the example plots we can see that both the Bspline basis (**38 splines**) and Fourier basis (**41 splines**) do a similar job of approximation of the original data. It can be noted that for both basis the new values perform a lot better than the old values.