

Supplement to the paper 'PU classification under Non-SCAR: clustering-assisted logistic model with oversampling enhancement'

Konrad Furmańczyk and Kacper Paczutkowski

Institute of Information Technology, Warsaw University of Life Sciences, Warsaw, Poland

This supplement contains 3 sections. In Section 1, tables for all classification metrics across 12 datasets and figures for the non-SCAR scheme with the executing time of considered algorithms are presented. Analogous results for the SCAR scheme are presented in Section 2. Section 3 contains the description of the non-SCAR labeling function. In all tables, we reported mean and standard deviations (in brackets). The best results are marked.

1 Non-SCAR Scheme

1.1 Tables for all classification metrics for 14 datasets.

q	naive acc	clust acc	strict lassclust acc	non strict lassclust acc	lasso joint acc	naive fl	clust fl	strict lassclust fl	non strict lassclust fl	lasso joint fl	naive auc	clust auc	strict lassclust auc	non strict lassclust auc	lasso joint auc
0.25	0.64 (0.16)	0.71 (0.16)	0.7 (0.15)	0.71 (0.16)	0.53 (0.15)	0.63 (0.18)	0.54 (0.25)	0.5 (0.25)	0.54 (0.25)	0.56 (0.14)	0.77 (0.13)	0.73 (0.17)	0.75 (0.16)	0.75 (0.16)	0.79 (0.15)
0.5	0.64 (0.16)	0.71 (0.16)	0.69 (0.15)	0.71 (0.16)	0.53 (0.15)	0.63 (0.18)	0.54 (0.24)	0.52 (0.23)	0.54 (0.25)	0.56 (0.14)	0.77 (0.13)	0.73 (0.17)	0.75 (0.16)	0.75 (0.16)	0.78 (0.15)
1	0.64 (0.16)	0.71 (0.16)	0.71 (0.16)	0.71 (0.16)	0.53 (0.15)	0.63 (0.18)	0.56 (0.24)	0.56 (0.24)	0.56 (0.24)	0.57 (0.15)	0.77 (0.13)	0.73 (0.17)	0.75 (0.16)	0.75 (0.16)	0.79 (0.15)

Table 1. Summary of Mean and Standard Deviation for Accuracy, F1, and AUC Metrics by q (Highest Mean in Bold); non-SCAR scheme

Table 2. Summary Statistics for df = adult; nonSCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.769 (0.005)	0.769 (0.005)	0.769 (0.005)	0.454 (0.005)	0.454 (0.005)	0.454 (0.005)	0.462 (0.006)	0.462 (0.006)	0.462 (0.006)
0.3	clust	0.588 (0.008)	0.761 (0.045)	0.763 (0.006)	0.646 (0.006)	0.724 (0.030)	0.621 (0.006)	0.279 (0.005)	0.502 (0.060)	0.513 (0.006)
0.3	strict-lassclust	0.589 (0.008)	0.761 (0.044)	0.763 (0.006)	0.646 (0.006)	0.724 (0.029)	0.621 (0.006)	0.279 (0.005)	0.502 (0.060)	0.513 (0.006)
0.3	non-strict-lassclust	0.589 (0.008)	0.761 (0.044)	0.763 (0.006)	0.646 (0.006)	0.724 (0.029)	0.621 (0.006)	0.279 (0.005)	0.502 (0.060)	0.513 (0.006)
0.3	lassojoint	0.751 (0.014)	0.751 (0.014)	0.751 (0.014)	0.363 (0.050)	0.363 (0.050)	0.363 (0.050)	0.427 (0.018)	0.427 (0.018)	0.427 (0.018)
0.5	naive	0.771 (0.003)	0.771 (0.003)	0.771 (0.003)	0.578 (0.003)	0.578 (0.003)	0.578 (0.003)	0.504 (0.006)	0.504 (0.006)	0.504 (0.006)
0.5	clust	0.704 (0.064)	0.780 (0.009)	0.766 (0.004)	0.711 (0.033)	0.731 (0.008)	0.595 (0.001)	0.361 (0.052)	0.522 (0.010)	0.511 (0.006)
0.5	strict-lassclust	0.704 (0.063)	0.780 (0.009)	0.766 (0.004)	0.711 (0.032)	0.732 (0.008)	0.595 (0.001)	0.357 (0.051)	0.522 (0.009)	0.511 (0.006)
0.5	non-strict-lassclust	0.704 (0.063)	0.780 (0.009)	0.766 (0.004)	0.711 (0.032)	0.732 (0.008)	0.595 (0.001)	0.357 (0.051)	0.522 (0.009)	0.511 (0.006)
0.5	lassojoint	0.766 (0.003)	0.766 (0.003)	0.766 (0.003)	0.586 (0.024)	0.586 (0.024)	0.586 (0.024)	0.501 (0.009)	0.501 (0.009)	0.501 (0.009)
0.8	naive	0.779 (0.000)	0.779 (0.000)	0.779 (0.000)	0.714 (0.002)	0.714 (0.002)	0.714 (0.002)	0.543 (0.002)	0.543 (0.002)	0.543 (0.002)
0.8	clust	0.787 (0.002)	0.774 (0.003)	0.770 (0.001)	0.749 (0.009)	0.675 (0.003)	0.455 (0.004)	0.530 (0.001)	0.536 (0.002)	0.469 (0.003)
0.8	strict-lassclust	0.787 (0.002)	0.775 (0.003)	0.770 (0.001)	0.749 (0.008)	0.675 (0.003)	0.454 (0.004)	0.530 (0.002)	0.536 (0.002)	0.469 (0.003)
0.8	non-strict-lassclust	0.787 (0.001)	0.775 (0.003)	0.770 (0.001)	0.749 (0.010)	0.675 (0.003)	0.454 (0.004)	0.529 (0.003)	0.536 (0.002)	0.469 (0.003)
0.8	lassojoint	0.775 (0.001)	0.775 (0.001)	0.775 (0.001)	0.735 (0.002)	0.735 (0.002)	0.735 (0.002)	0.531 (0.002)	0.531 (0.002)	0.531 (0.002)

Table 3. Summary Statistics for df = artif; nonSCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.764 (0.017)	0.764 (0.017)	0.764 (0.017)	0.585 (0.018)	0.585 (0.018)	0.585 (0.018)	0.694 (0.016)	0.694 (0.016)	0.694 (0.016)
0.3	clust	0.623 (0.081)	0.580 (0.068)	0.552 (0.039)	0.487 (0.104)	0.515 (0.074)	0.518 (0.036)	0.452 (0.119)	0.495 (0.081)	0.591 (0.037)
0.3	strict-lassclust	0.617 (0.080)	0.577 (0.066)	0.552 (0.039)	0.491 (0.095)	0.514 (0.067)	0.520 (0.034)	0.430 (0.154)	0.451 (0.153)	0.580 (0.087)
0.3	non-strict-lassclust	0.619 (0.080)	0.579 (0.067)	0.553 (0.040)	0.489 (0.100)	0.516 (0.073)	0.522 (0.036)	0.453 (0.113)	0.498 (0.081)	0.589 (0.040)
0.3	lassojoint	0.721 (0.025)	0.723 (0.023)	0.725 (0.021)	0.543 (0.052)	0.547 (0.050)	0.550 (0.050)	0.675 (0.024)	0.676 (0.023)	0.677 (0.023)
0.5	naive	0.850 (0.011)	0.850 (0.011)	0.850 (0.011)	0.674 (0.004)	0.674 (0.004)	0.674 (0.004)	0.741 (0.006)	0.741 (0.006)	0.741 (0.006)
0.5	clust	0.660 (0.014)	0.598 (0.102)	0.564 (0.012)	0.610 (0.005)	0.507 (0.105)	0.526 (0.003)	0.566 (0.012)	0.494 (0.108)	0.633 (0.008)
0.5	strict-lassclust	0.640 (0.005)	0.573 (0.095)	0.573 (0.013)	0.587 (0.007)	0.517 (0.070)	0.504 (0.014)	0.556 (0.066)	0.647 (0.041)	0.648 (0.009)
0.5	non-strict-lassclust	0.650 (0.024)	0.594 (0.104)	0.566 (0.008)	0.608 (0.000)	0.519 (0.093)	0.524 (0.005)	0.555 (0.036)	0.432 (0.124)	0.607 (0.021)
0.5	lassojoint	0.843 (0.015)	0.843 (0.015)	0.843 (0.015)	0.675 (0.006)	0.676 (0.005)	0.676 (0.006)	0.738 (0.006)	0.739 (0.006)	0.739 (0.006)
0.8	naive	0.942 (0.000)	0.942 (0.000)	0.942 (0.000)	0.820 (0.000)	0.820 (0.000)	0.820 (0.000)	0.841 (0.000)	0.841 (0.000)	0.841 (0.000)
0.8	clust	0.707 (0.000)	0.674 (0.000)	0.641 (0.000)	0.660 (0.000)	0.625 (0.000)	0.525 (0.000)	0.669 (0.000)	0.639 (0.000)	0.675 (0.000)
0.8	strict-lassclust	0.706 (0.000)	0.725 (0.000)	0.653 (0.000)	0.658 (0.000)	0.657 (0.000)	0.552 (0.000)	0.667 (0.000)	0.672 (0.000)	0.679 (0.000)
0.8	non-strict-lassclust	0.706 (0.000)	0.658 (0.000)	0.637 (0.000)	0.658 (0.000)	0.618 (0.000)	0.513 (0.000)	0.667 (0.000)	0.636 (0.000)	0.670 (0.000)
0.8	lassojoint	0.847 (0.000)	0.877 (0.000)	0.847 (0.000)	0.513 (0.000)	0.757 (0.000)	0.513 (0.000)	0.671 (0.000)	0.793 (0.000)	0.671 (0.000)

Table 4. Summary Statistics for df = bank-marketing; nonSCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.560 (0.028)	0.560 (0.028)	0.560 (0.028)	0.300 (0.010)	0.300 (0.010)	0.300 (0.010)	0.160 (0.012)	0.160 (0.012)	0.160 (0.012)
0.3	clust	0.710 (0.020)	0.710 (0.020)	0.702 (0.020)	0.282 (0.011)	0.282 (0.011)	0.280 (0.011)	0.084 (0.011)	0.084 (0.011)	0.088 (0.011)
0.3	strict-lassclust	0.710 (0.021)	0.710 (0.021)	0.710 (0.021)	0.282 (0.011)	0.282 (0.011)	0.282 (0.011)	0.084 (0.011)	0.084 (0.011)	0.084 (0.011)
0.3	non-strict-lassclust	0.710 (0.021)	0.710 (0.021)	0.710 (0.021)	0.282 (0.011)	0.282 (0.011)	0.282 (0.011)	0.084 (0.011)	0.084 (0.011)	0.084 (0.011)
0.3	lassojoint	0.600 (0.041)	0.601 (0.040)	0.604 (0.041)	0.210 (0.089)	0.207 (0.090)	0.195 (0.089)	0.180 (0.036)	0.181 (0.033)	0.182 (0.036)
0.5	naive	0.561 (0.014)	0.561 (0.014)	0.561 (0.014)	0.380 (0.005)	0.380 (0.005)	0.380 (0.005)	0.126 (0.008)	0.126 (0.008)	0.126 (0.008)
0.5	clust	0.734 (0.021)	0.734 (0.021)	0.433 (0.196)	0.240 (0.033)	0.240 (0.032)	0.119 (0.018)	0.072 (0.010)	0.072 (0.010)	0.106 (0.018)
0.5	strict-lassclust	0.791 (0.016)	0.791 (0.016)	0.776 (0.018)	0.290 (0.005)	0.290 (0.005)	0.290 (0.005)	0.073 (0.006)	0.073 (0.006)	0.073 (0.006)
0.5	non-strict-lassclust	0.791 (0.016)	0.791 (0.016)	0.776 (0.018)	0.290 (0.005)	0.290 (0.005)	0.290 (0.005)	0.073 (0.006)	0.073 (0.006)	0.073 (0.006)
0.5	lassojoint	0.612 (0.015)	0.612 (0.015)	0.612 (0.015)	0.533 (0.008)	0.533 (0.008)	0.533 (0.008)	0.126 (0.013)	0.126 (0.013)	0.126 (0.013)
0.8	naive	0.648 (0.013)	0.648 (0.013)	0.648 (0.013)	0.774 (0.005)	0.774 (0.005)	0.774 (0.005)	0.261 (0.009)	0.261 (0.009)	0.261 (0.009)
0.8	clust	0.673 (0.004)	0.673 (0.004)	0.461 (0.009)	0.300 (0.005)	0.300 (0.005)	0.107 (0.001)	0.094 (0.004)	0.094 (0.004)	0.168 (0.007)
0.8	strict-lassclust	0.665 (0.003)	0.665 (0.003)	0.672 (0.003)	0.300 (0.005)	0.300 (0.005)	0.300 (0.005)	0.094 (0.003)	0.094 (0.003)	0.094 (0.003)
0.8	non-strict-lassclust	0.665 (0.003)	0.665 (0.003)	0.672 (0.003)	0.300 (0.005)	0.300 (0.005)	0.300 (0.005)	0.094 (0.003)	0.094 (0.003)	0.094 (0.003)
0.8	lassojoint	0.547 (0.013)	0.547 (0.013)	0.547 (0.013)	0.773 (0.004)	0.773 (0.004)	0.773 (0.004)	0.200 (0.003)	0.200 (0.003)	0.200 (0.003)

Table 5. Summary Statistics for df = banknote; nonSCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.810 (0.024)	0.810 (0.024)	0.810 (0.024)	0.470 (0.022)	0.470 (0.022)	0.470 (0.022)	0.620 (0.021)	0.620 (0.021)	0.620 (0.021)
0.3	clust	0.778 (0.047)	0.753 (0.042)	0.728 (0.063)	0.315 (0.038)	0.333 (0.041)	0.324 (0.039)	0.357 (0.039)	0.348 (0.028)	0.437 (0.033)
0.3	strict-lassclust	0.768 (0.044)	0.741 (0.035)	0.722 (0.036)	0.326 (0.034)	0.347 (0.032)	0.326 (0.045)	0.361 (0.037)	0.353 (0.026)	0.426 (0.026)
0.3	non-strict-lassclust	0.767 (0.044)	0.741 (0.036)	0.723 (0.037)	0.328 (0.032)	0.347 (0.033)	0.328 (0.045)	0.360 (0.038)	0.353 (0.027)	0.428 (0.033)
0.3	lassojoint	0.605 (0.157)	0.610 (0.164)	0.603 (0.160)	0.417 (0.092)	0.408 (0.098)	0.413 (0.094)	0.474 (0.155)	0.461 (0.165)	0.469 (0.158)
0.5	naive	0.915 (0.000)	0.915 (0.000)	0.915 (0.000)	0.591 (0.009)	0.591 (0.009)	0.591 (0.009)	0.679 (0.007)	0.679 (0.007)	0.679 (0.007)
0.5	clust	0.906 (0.001)	0.905 (0.001)	0.803 (0.002)	0.280 (0.005)	0.276 (0.005)	0.341 (0.007)	0.407 (0.007)	0.400 (0.006)	0.495 (0.010)
0.5	strict-lassclust	0.902 (0.001)	0.901 (0.001)	0.801 (0.001)	0.280 (0.005)	0.276 (0.002)	0.341 (0.007)	0.407 (0.007)	0.400 (0.005)	0.495 (0.010)
0.5	non-strict-lassclust	0.902 (0.001)	0.901 (0.001)	0.801 (0.001)	0.280 (0.005)	0.276 (0.002)	0.341 (0.007)	0.407 (0.007)	0.400 (0.005)	0.495 (0.010)
0.5	lassojoint	0.926 (0.000)	0.926 (0.000)	0.926 (0.000)	0.509 (0.006)	0.509 (0.006)	0.509 (0.006)	0.645 (0.006)	0.645 (0.006)	0.645 (0.006)
0.8	naive	1.000 (0.000)	1.000 (0.000)	1.000 (0.000)	0.990 (0.000)	0.990 (0.000)	0.990 (0.000)	0.988 (0.000)	0.988 (0.000)	0.988 (0.000)
0.8	clust	0.610 (0.000)	0.594 (0.000)	0.571 (0.000)	0.628 (0.000)	0.616 (0.000)	0.506 (0.000)	0.481 (0.000)	0.506 (0.000)	0.609 (0.000)
0.8	strict-lassclust	0.659 (0.000)	0.626 (0.000)	0.586 (0.000)	0.637 (0.000)	0.582 (0.000)	0.482 (0.000)	0.488 (0.000)	0.494 (0.000)	0.600 (0.000)
0.8	non-strict-lassclust	0.659 (0.000)	0.626 (0.000)	0.586 (0.000)	0.637 (0.000)	0.582 (0.000)	0.482 (0.000)	0.488 (0.000)	0.494 (0.000)	0.600 (0.000)
0.8	lassojoint	0.985 (0.000)	0.985 (0.000)	0.985 (0.000)	0.839 (0.000)	0.839 (0.000)	0.839 (0.000)	0.835 (0.000)	0.835 (0.000)	0.835 (0.000)

Table 6. Summary Statistics for df = breastc; nonSCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.924 (0.013)	0.924 (0.013)	0.924 (0.013)	0.464 (0.032)	0.464 (0.032)	0.464 (0.032)	0.565 (0.031)	0.565 (0.031)	0.565 (0.031)
0.3	clust	0.965 (0.015)	0.964 (0.015)	0.957 (0.021)	0.957 (0.015)	0.958 (0.015)	0.953 (0.018)	0.937 (0.022)	0.939 (0.022)	0.930 (0.029)
0.3	strict-lassclust	0.994 (0.003)	0.994 (0.003)	0.994 (0.003)	0.960 (0.014)	0.960 (0.014)	0.959 (0.015)	0.941 (0.020)	0.941 (0.021)	0.941 (0.022)
0.3	non-strict-lassclust	0.994 (0.003)	0.994 (0.003)	0.994 (0.003)	0.960 (0.014)	0.960 (0.014)	0.959 (0.015)	0.942 (0.020)	0.942 (0.021)	0.941 (0.022)
0.3	lassojoint	0.987 (0.007)	0.987 (0.007)	0.987 (0.007)	0.365 (0.060)	0.365 (0.059)	0.364 (0.059)	0.524 (0.039)	0.524 (0.039)	0.523 (0.039)
0.5	naive	0.938 (0.002)	0.938 (0.002)	0.938 (0.002)	0.753 (0.014)	0.753 (0.014)	0.753 (0.014)	0.744 (0.014)	0.744 (0.014)	0.744 (0.014)
0.5	clust	0.968 (0.003)	0.968 (0.004)	0.976 (0.007)	0.962 (0.002)	0.962 (0.002)	0.970 (0.004)	0.945 (0.003)	0.945 (0.004)	0.956 (0.006)
0.5	strict-lassclust	0.997 (0.000)	0.997 (0.000)	0.998 (0.000)	0.967 (0.002)	0.967 (0.002)	0.975 (0.001)	0.953 (0.003)	0.953 (0.003)	0.966 (0.002)
0.5	non-strict-lassclust	0.997 (0.000)	0.997 (0.000)	0.998 (0.000)	0.967 (0.002)	0.967 (0.002)	0.975 (0.001)	0.953 (0.003)	0.953 (0.003)	0.966 (0.002)
0.5	lassojoint	0.998 (0.001)	0.998 (0.001)	0.998 (0.001)	0.359 (0.006)	0.359 (0.006)	0.359 (0.006)	0.528 (0.006)	0.528 (0.006)	0.528 (0.006)
0.8	naive	0.990 (0.001)	0.990 (0.001)	0.990 (0.001)	0.935 (0.003)	0.935 (0.003)	0.935 (0.003)	0.910 (0.004)	0.910 (0.004)	0.910 (0.004)
0.8	clust	0.948 (0.004)	0.956 (0.005)	0.969 (0.002)	0.960 (0.003)	0.962 (0.003)	0.969 (0.007)	0.936 (0.004)	0.939 (0.005)	0.952 (0.011)
0.8	strict-lassclust	0.997 (0.001)	0.997 (0.001)	0.996 (0.000)	0.961 (0.000)	0.961 (0.002)	0.960 (0.006)	0.937 (0.001)	0.937 (0.004)	0.937 (0.009)
0.8	non-strict-lassclust	0.997 (0.001)	0.997 (0.000)	0.996 (0.000)	0.961 (0.000)	0.962 (0.003)	0.960 (0.006)	0.937 (0.001)	0.939 (0.006)	0.937 (0.009)
0.8	lassojoint	0.993 (0.000)	0.993 (0.001)	0.993 (0.000)	0.854 (0.034)	0.854 (0.033)	0.854 (0.034)	0.819 (0.041)	0.818 (0.039)	0.819 (0.041)

Table 7. Summary Statistics for df = credit-a; nonSCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.714 (0.035)	0.714 (0.035)	0.714 (0.035)	0.542 (0.027)	0.542 (0.027)	0.542 (0.027)	0.647 (0.028)	0.647 (0.028)	0.647 (0.028)
0.3	clust	0.609 (0.028)	0.613 (0.027)	0.622 (0.027)	0.633 (0.031)	0.636 (0.031)	0.644 (0.029)	0.376 (0.062)	0.387 (0.062)	0.413 (0.056)
0.3	strict-lassclust	0.714 (0.042)	0.717 (0.042)	0.715 (0.040)	0.631 (0.029)	0.634 (0.030)	0.643 (0.030)	0.373 (0.057)	0.382 (0.059)	0.410 (0.057)
0.3	non-strict-lassclust	0.714 (0.042)	0.718 (0.042)	0.715 (0.040)	0.631 (0.029)	0.634 (0.030)	0.643 (0.030)	0.373 (0.057)	0.383 (0.059)	0.410 (0.057)
0.3	lassojoint	0.679 (0.035)	0.680 (0.035)	0.679 (0.035)	0.529 (0.045)	0.529 (0.045)	0.529 (0.045)	0.639 (0.028)	0.639 (0.028)	0.639 (0.028)
0.5	naive	0.721 (0.008)	0.721 (0.008)	0.721 (0.008)	0.574 (0.006)	0.574 (0.006)	0.574 (0.006)	0.639 (0.006)	0.639 (0.006)	0.639 (0.006)
0.5	clust	0.585 (0.004)	0.586 (0.005)	0.589 (0.007)	0.605 (0.002)	0.607 (0.005)	0.616 (0.004)	0.297 (0.011)	0.311 (0.018)	0.353 (0.017)
0.5	strict-lassclust	0.725 (0.006)	0.731 (0.005)	0.728 (0.008)	0.605 (0.002)	0.602 (0.006)	0.622 (0.003)	0.298 (0.014)	0.297 (0.022)	0.366 (0.013)
0.5	non-strict-lassclust	0.725 (0.006)	0.731 (0.005)	0.728 (0.008)	0.605 (0.002)	0.602 (0.006)	0.622 (0.003)	0.298 (0.014)	0.297 (0.022)	0.366 (0.013)
0.5	lassojoint	0.705 (0.010)	0.705 (0.010)	0.705 (0.010)	0.609 (0.026)	0.609 (0.026)	0.609 (0.026)	0.653 (0.005)	0.653 (0.005)	0.653 (0.005)
0.8	naive	0.827 (0.000)	0.827 (0.000)	0.827 (0.000)	0.687 (0.000)	0.687 (0.000)	0.687 (0.000)	0.687 (0.000)	0.687 (0.000)	0.687 (0.000)
0.8	clust	0.634 (0.000)	0.642 (0.000)	0.649 (0.000)	0.687 (0.000)	0.692 (0.000)	0.703 (0.000)	0.396 (0.000)	0.412 (0.000)	0.463 (0.000)
0.8	strict-lassclust	0.786 (0.000)	0.778 (0.000)	0.794 (0.000)	0.692 (0.000)	0.692 (0.000)	0.703 (0.000)	0.412 (0.000)	0.423 (0.000)	0.453 (0.000)
0.8	non-strict-lassclust	0.786 (0.000)	0.778 (0.000)	0.794 (0.000)	0.692 (0.000)	0.692 (0.000)	0.703 (0.000)	0.412 (0.000)	0.423 (0.000)	0.453 (0.000)
0.8	lassojoint	0.824 (0.000)	0.824 (0.000)	0.824 (0.000)	0.738 (0.000)	0.738 (0.000)	0.738 (0.000)	0.721 (0.000)	0.721 (0.000)	0.721 (0.000)

Table 8. Summary Statistics for df = credit-g; nonSCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.579 (0.036)	0.579 (0.036)	0.579 (0.036)	0.340 (0.023)	0.340 (0.023)	0.340 (0.023)	0.420 (0.028)	0.420 (0.028)	0.420 (0.028)
0.3	clust	0.566 (0.044)	0.568 (0.048)	0.567 (0.044)	0.644 (0.045)	0.653 (0.035)	0.600 (0.034)	0.328 (0.050)	0.339 (0.048)	0.387 (0.039)
0.3	strict-lassclust	0.570 (0.040)	0.572 (0.046)	0.572 (0.042)	0.665 (0.031)	0.666 (0.031)	0.617 (0.033)	0.324 (0.085)	0.338 (0.078)	0.380 (0.040)
0.3	non-strict-lassclust	0.568 (0.045)	0.574 (0.046)	0.571 (0.043)	0.638 (0.070)	0.644 (0.073)	0.605 (0.051)	0.336 (0.047)	0.344 (0.046)	0.383 (0.039)
0.3	lassojoint	0.602 (0.033)	0.602 (0.033)	0.602 (0.034)	0.394 (0.026)	0.394 (0.026)	0.394 (0.027)	0.344 (0.033)	0.344 (0.033)	0.344 (0.033)
0.5	naive	0.569 (0.024)	0.569 (0.024)	0.569 (0.024)	0.408 (0.002)	0.408 (0.002)	0.408 (0.002)	0.371 (0.024)	0.371 (0.024)	0.371 (0.024)
0.5	clust	0.549 (0.025)	0.530 (0.020)	0.519 (0.021)	0.623 (0.042)	0.663 (0.014)	0.492 (0.046)	0.311 (0.053)	0.336 (0.007)	0.367 (0.006)
0.5	strict-lassclust	0.548 (0.026)	0.541 (0.007)	0.522 (0.019)	0.647 (0.050)	0.665 (0.006)	0.498 (0.044)	0.307 (0.047)	0.257 (0.102)	0.370 (0.005)
0.5	non-strict-lassclust	0.529 (0.035)	0.539 (0.017)	0.522 (0.019)	0.547 (0.103)	0.643 (0.039)	0.498 (0.044)	0.321 (0.062)	0.323 (0.012)	0.370 (0.005)
0.5	lassojoint	0.624 (0.012)	0.624 (0.012)	0.625 (0.013)	0.437 (0.007)	0.437 (0.007)	0.437 (0.007)	0.255 (0.030)	0.255 (0.030)	0.255 (0.030)
0.8	naive	0.524 (0.009)	0.524 (0.009)	0.524 (0.009)	0.579 (0.005)	0.579 (0.005)	0.579 (0.005)	0.280 (0.018)	0.280 (0.018)	0.280 (0.018)
0.8	clust	0.559 (0.065)	0.519 (0.014)	0.528 (0.029)	0.585 (0.068)	0.616 (0.015)	0.363 (0.013)	0.283 (0.063)	0.317 (0.032)	0.423 (0.010)
0.8	strict-lassclust	0.572 (0.073)	0.521 (0.005)	0.540 (0.018)	0.606 (0.071)	0.635 (0.019)	0.364 (0.008)	0.195 (0.058)	0.320 (0.036)	0.420 (0.008)
0.8	non-strict-lassclust	0.555 (0.061)	0.521 (0.010)	0.540 (0.018)	0.576 (0.061)	0.618 (0.016)	0.364 (0.008)	0.278 (0.059)	0.313 (0.029)	0.420 (0.008)
0.8	lassojoint	0.643 (0.012)	0.643 (0.012)	0.643 (0.012)	0.610 (0.000)	0.610 (0.000)	0.610 (0.000)	0.033 (0.000)	0.033 (0.000)	0.033 (0.000)

Table 9. Summary Statistics for df = dhfr; nonSCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.538 (0.048)	0.538 (0.048)	0.538 (0.048)	0.466 (0.051)	0.466 (0.051)	0.466 (0.051)	0.481 (0.059)	0.481 (0.059)	0.481 (0.059)
0.3	clust	0.683 (0.111)	0.657 (0.084)	0.567 (0.120)	0.303 (0.082)	0.340 (0.074)	0.357 (0.096)	0.323 (0.087)	0.341 (0.091)	0.422 (0.068)
0.3	strict-lassclust	0.635 (0.101)	0.646 (0.094)	0.590 (0.079)	0.492 (0.138)	0.492 (0.113)	0.446 (0.065)	0.538 (0.243)	0.426 (0.244)	0.381 (0.088)
0.3	non-strict-lassclust	0.624 (0.090)	0.612 (0.056)	0.582 (0.070)	0.442 (0.111)	0.483 (0.099)	0.449 (0.054)	0.372 (0.175)	0.306 (0.184)	0.392 (0.079)
0.3	lassojoint	0.662 (0.061)	0.648 (0.049)	0.660 (0.064)	0.453 (0.087)	0.490 (0.048)	0.504 (0.085)	0.552 (0.041)	0.556 (0.033)	0.562 (0.025)
0.5	naive	0.728 (0.000)	0.728 (0.000)	0.728 (0.000)	0.629 (0.000)	0.629 (0.000)	0.629 (0.000)	0.667 (0.000)	0.667 (0.000)	0.667 (0.000)
0.5	clust	0.518 (0.000)	0.518 (0.000)	0.552 (0.000)	0.433 (0.000)	0.433 (0.000)	0.454 (0.000)	0.396 (0.000)	0.396 (0.000)	0.505 (0.000)
0.5	strict-lassclust	0.606 (0.000)	0.615 (0.000)	0.594 (0.000)	0.515 (0.000)	0.454 (0.000)	0.371 (0.000)	0.373 (0.000)	0.361 (0.000)	0.396 (0.000)
0.5	non-strict-lassclust	0.583 (0.000)	0.587 (0.000)	0.592 (0.000)	0.392 (0.000)	0.485 (0.000)	0.351 (0.000)	0.366 (0.000)	0.405 (0.000)	0.364 (0.000)
0.5	lassojoint	0.706 (0.000)	0.704 (0.000)	0.704 (0.000)	0.588 (0.000)	0.464 (0.000)	0.464 (0.000)	0.661 (0.000)	0.612 (0.000)	0.612 (0.000)
0.8	naive	0.653 (0.000)	0.653 (0.000)	0.653 (0.000)	0.608 (0.000)	0.608 (0.000)	0.608 (0.000)	0.578 (0.000)	0.578 (0.000)	0.578 (0.000)
0.8	clust	0.500 (0.000)	0.547 (0.000)	0.512 (0.000)	0.474 (0.000)	0.433 (0.000)	0.412 (0.000)	0.495 (0.000)	0.409 (0.000)	0.571 (0.000)
0.8	strict-lassclust	0.507 (0.000)	0.492 (0.000)	0.610 (0.000)	0.433 (0.000)	0.433 (0.000)	0.454 (0.000)	0.604 (0.000)	0.604 (0.000)	0.583 (0.000)
0.8	non-strict-lassclust	0.622 (0.000)	0.603 (0.000)	0.581 (0.000)	0.629 (0.000)	0.608 (0.000)	0.423 (0.000)	0.538 (0.000)	0.387 (0.000)	0.582 (0.000)
0.8	lassojoint	0.839 (0.000)	0.845 (0.000)	0.845 (0.000)	0.691 (0.000)	0.680 (0.000)	0.732 (0.000)	0.706 (0.000)	0.710 (0.000)	0.740 (0.000)

Table 10. Summary Statistics for df = diabetes; nonSCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.705 (0.028)	0.705 (0.028)	0.705 (0.028)	0.500 (0.029)	0.500 (0.029)	0.500 (0.029)	0.566 (0.033)	0.566 (0.033)	0.566 (0.033)
0.3	clust	0.677 (0.073)	0.690 (0.055)	0.690 (0.055)	0.665 (0.046)	0.676 (0.031)	0.676 (0.025)	0.527 (0.054)	0.545 (0.047)	0.564 (0.047)
0.3	strict-lassclust	0.686 (0.072)	0.711 (0.046)	0.722 (0.040)	0.667 (0.042)	0.680 (0.032)	0.677 (0.026)	0.485 (0.113)	0.533 (0.065)	0.562 (0.051)
0.3	non-strict-lassclust	0.691 (0.068)	0.712 (0.048)	0.723 (0.041)	0.661 (0.053)	0.678 (0.032)	0.675 (0.028)	0.532 (0.057)	0.550 (0.048)	0.567 (0.049)
0.3	lassojoint	0.717 (0.027)	0.717 (0.027)	0.717 (0.027)	0.379 (0.032)	0.379 (0.032)	0.379 (0.032)	0.526 (0.031)	0.526 (0.031)	0.526 (0.031)
0.5	naive	0.700 (0.012)	0.700 (0.012)	0.700 (0.012)	0.595 (0.013)	0.595 (0.013)	0.595 (0.013)	0.604 (0.008)	0.604 (0.008)	0.604 (0.008)
0.5	clust	0.702 (0.052)	0.660 (0.051)	0.628 (0.013)	0.659 (0.065)	0.664 (0.026)	0.656 (0.007)	0.476 (0.077)	0.497 (0.054)	0.521 (0.013)
0.5	strict-lassclust	0.687 (0.046)	0.708 (0.035)	0.692 (0.012)	0.664 (0.058)	0.669 (0.023)	0.653 (0.007)	0.367 (0.192)	0.380 (0.203)	0.522 (0.036)
0.5	non-strict-lassclust	0.696 (0.048)	0.702 (0.017)	0.690 (0.011)	0.654 (0.059)	0.658 (0.020)	0.655 (0.009)	0.496 (0.073)	0.520 (0.049)	0.521 (0.034)
0.5	lassojoint	0.694 (0.010)	0.693 (0.011)	0.693 (0.009)	0.453 (0.122)	0.453 (0.122)	0.454 (0.120)	0.547 (0.053)	0.546 (0.052)	0.548 (0.053)
0.8	naive	0.758 (0.010)	0.758 (0.010)	0.758 (0.010)	0.697 (0.015)	0.697 (0.015)	0.697 (0.015)	0.627 (0.020)	0.627 (0.020)	0.627 (0.020)
0.8	clust	0.604 (0.031)	0.658 (0.013)	0.695 (0.020)	0.630 (0.019)	0.643 (0.010)	0.613 (0.043)	0.427 (0.031)	0.467 (0.020)	0.575 (0.004)
0.8	strict-lassclust	0.561 (0.014)	0.667 (0.016)	0.693 (0.020)	0.625 (0.022)	0.651 (0.017)	0.610 (0.041)	0.246 (0.284)	0.487 (0.031)	0.566 (0.012)
0.8	non-strict-lassclust	0.618 (0.037)	0.657 (0.015)	0.696 (0.020)	0.622 (0.017)	0.645 (0.011)	0.607 (0.040)	0.485 (0.040)	0.471 (0.021)	0.566 (0.011)
0.8	lassojoint	0.751 (0.004)	0.754 (0.006)	0.759 (0.010)	0.690 (0.010)	0.695 (0.007)	0.696 (0.005)	0.613 (0.014)	0.621 (0.010)	0.620 (0.010)

Table 11. Summary Statistics for df = heart-c; nonSCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.584 (0.056)	0.584 (0.056)	0.584 (0.056)	0.494 (0.042)	0.494 (0.042)	0.494 (0.042)	0.622 (0.040)	0.622 (0.040)	0.622 (0.040)
0.3	clust	0.699 (0.046)	0.697 (0.047)	0.693 (0.047)	0.658 (0.044)	0.654 (0.044)	0.627 (0.045)	0.630 (0.053)	0.637 (0.051)	0.666 (0.045)
0.3	strict-lassclust	0.700 (0.046)	0.697 (0.047)	0.692 (0.048)	0.662 (0.044)	0.657 (0.043)	0.624 (0.047)	0.634 (0.053)	0.640 (0.049)	0.665 (0.044)
0.3	non-strict-lassclust	0.700 (0.046)	0.697 (0.047)	0.692 (0.048)	0.662 (0.045)	0.657 (0.043)	0.624 (0.047)	0.635 (0.053)	0.641 (0.048)	0.665 (0.044)
0.3	lassojoint	0.564 (0.063)	0.566 (0.064)	0.565 (0.062)	0.468 (0.047)	0.467 (0.047)	0.467 (0.046)	0.540 (0.066)	0.539 (0.065)	0.538 (0.065)
0.5	naive	0.570 (0.000)	0.570 (0.000)	0.570 (0.000)	0.578 (0.000)	0.578 (0.000)	0.578 (0.000)	0.678 (0.000)	0.678 (0.000)	0.678 (0.000)
0.5	clust	0.642 (0.000)	0.645 (0.000)	0.640 (0.000)	0.611 (0.000)	0.633 (0.000)	0.600 (0.000)	0.646 (0.000)	0.673 (0.000)	0.700 (0.000)
0.5	strict-lassclust	0.641 (0.000)	0.642 (0.000)	0.635 (0.000)	0.600 (0.000)	0.633 (0.000)	0.600 (0.000)	0.633 (0.000)	0.673 (0.000)	0.700 (0.000)
0.5	non-strict-lassclust	0.641 (0.000)	0.642 (0.000)	0.635 (0.000)	0.600 (0.000)	0.633 (0.000)	0.600 (0.000)	0.633 (0.000)	0.673 (0.000)	0.700 (0.000)
0.5	lassojoint	0.632 (0.000)	0.632 (0.000)	0.594 (0.000)	0.444 (0.000)	0.444 (0.000)	0.589 (0.000)	0.479 (0.000)	0.479 (0.000)	0.694 (0.000)
0.8	naive	0.754 (0.000)	0.754 (0.000)	0.754 (0.000)	0.667 (0.000)	0.667 (0.000)	0.667 (0.000)	0.732 (0.000)	0.732 (0.000)	0.732 (0.000)
0.8	clust	0.683 (0.000)	0.701 (0.000)	0.685 (0.000)	0.689 (0.000)	0.678 (0.000)	0.567 (0.000)	0.745 (0.000)	0.739 (0.000)	0.715 (0.000)
0.8	strict-lassclust	0.688 (0.000)	0.683 (0.000)	0.686 (0.000)	0.678 (0.000)	0.656 (0.000)	0.567 (0.000)	0.743 (0.000)	0.739 (0.000)	0.715 (0.000)
0.8	non-strict-lassclust	0.693 (0.000)	0.704 (0.000)	0.686 (0.000)	0.678 (0.000)	0.667 (0.000)	0.567 (0.000)	0.734 (0.000)	0.727 (0.000)	0.715 (0.000)
0.8	lassojoint	0.759 (0.000)	0.759 (0.000)	0.759 (0.000)	0.700 (0.000)	0.700 (0.000)	0.700 (0.000)	0.773 (0.000)	0.773 (0.000)	0.773 (0.000)

Table 12. Summary Statistics for df = qsar-biodeg; nonSCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.633 (0.026)	0.633 (0.026)	0.633 (0.026)	0.484 (0.027)	0.484 (0.027)	0.484 (0.027)	0.547 (0.030)	0.547 (0.030)	0.547 (0.030)
0.3	clust	0.765 (0.028)	0.766 (0.027)	0.761 (0.034)	0.670 (0.031)	0.666 (0.032)	0.575 (0.050)	0.626 (0.033)	0.625 (0.032)	0.582 (0.055)
0.3	strict-lassclust	0.760 (0.043)	0.762 (0.031)	0.763 (0.035)	0.662 (0.079)	0.668 (0.071)	0.623 (0.091)	0.477 (0.195)	0.481 (0.191)	0.515 (0.153)
0.3	non-strict-lassclust	0.740 (0.055)	0.741 (0.055)	0.739 (0.052)	0.542 (0.130)	0.535 (0.129)	0.490 (0.093)	0.566 (0.082)	0.573 (0.055)	0.548 (0.078)
0.3	lassojoint	0.662 (0.033)	0.664 (0.030)	0.662 (0.033)	0.436 (0.031)	0.436 (0.031)	0.437 (0.032)	0.534 (0.028)	0.535 (0.029)	0.535 (0.028)
0.5	naive	0.712 (0.001)	0.712 (0.001)	0.712 (0.001)	0.561 (0.013)	0.561 (0.013)	0.561 (0.013)	0.560 (0.002)	0.560 (0.002)	0.560 (0.002)
0.5	clust	0.759 (0.008)	0.761 (0.010)	0.763 (0.009)	0.685 (0.012)	0.674 (0.029)	0.548 (0.012)	0.638 (0.019)	0.629 (0.029)	0.572 (0.013)
0.5	strict-lassclust	0.745 (0.015)	0.759 (0.009)	0.763 (0.010)	0.612 (0.096)	0.599 (0.092)	0.536 (0.039)	0.586 (0.044)	0.597 (0.044)	0.571 (0.022)
0.5	non-strict-lassclust	0.758 (0.002)	0.751 (0.019)	0.760 (0.008)	0.631 (0.045)	0.667 (0.009)	0.475 (0.040)	0.592 (0.009)	0.612 (0.020)	0.550 (0.008)
0.5	lassojoint	0.699 (0.015)	0.699 (0.015)	0.707 (0.005)	0.434 (0.071)	0.440 (0.078)	0.463 (0.038)	0.529 (0.015)	0.531 (0.017)	0.540 (0.003)
0.8	naive	0.815 (0.007)	0.815 (0.007)	0.815 (0.007)	0.734 (0.000)	0.734 (0.000)	0.734 (0.000)	0.667 (0.000)	0.667 (0.000)	0.667 (0.000)
0.8	clust	0.801 (0.009)	0.791 (0.010)	0.779 (0.009)	0.739 (0.029)	0.700 (0.021)	0.513 (0.007)	0.658 (0.030)	0.644 (0.023)	0.566 (0.002)
0.8	strict-lassclust	0.774 (0.027)	0.774 (0.007)	0.765 (0.009)	0.696 (0.052)	0.659 (0.007)	0.444 (0.019)	0.623 (0.010)	0.630 (0.001)	0.538 (0.007)
0.8	non-strict-lassclust	0.778 (0.002)	0.786 (0.011)	0.765 (0.010)	0.677 (0.040)	0.698 (0.019)	0.426 (0.008)	0.631 (0.017)	0.625 (0.033)	0.531 (0.003)
0.8	lassojoint	0.746 (0.003)	0.742 (0.003)	0.750 (0.011)	0.586 (0.021)	0.583 (0.021)	0.592 (0.030)	0.574 (0.010)	0.571 (0.008)	0.574 (0.011)

Table 13. Summary Statistics for df = spambase; nonSCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.685 (0.010)	0.685 (0.010)	0.685 (0.010)	0.552 (0.016)	0.552 (0.016)	0.552 (0.016)	0.632 (0.013)	0.632 (0.013)	0.632 (0.013)
0.3	clust	0.831 (0.066)	0.844 (0.074)	0.876 (0.066)	0.604 (0.187)	0.693 (0.179)	0.814 (0.113)	0.661 (0.099)	0.710 (0.092)	0.777 (0.063)
0.3	strict-lassclust	0.847 (0.051)	0.865 (0.053)	0.905 (0.024)	0.593 (0.168)	0.665 (0.170)	0.802 (0.135)	0.659 (0.088)	0.694 (0.089)	0.772 (0.073)
0.3	non-strict-lassclust	0.849 (0.047)	0.868 (0.043)	0.904 (0.026)	0.579 (0.158)	0.653 (0.168)	0.801 (0.134)	0.653 (0.083)	0.692 (0.089)	0.772 (0.073)
0.3	lassojoint	0.806 (0.015)	0.806 (0.015)	0.806 (0.016)	0.433 (0.017)	0.433 (0.017)	0.433 (0.017)	0.579 (0.013)	0.579 (0.013)	0.579 (0.013)
0.5	naive	0.786 (0.004)	0.786 (0.004)	0.786 (0.004)	0.671 (0.004)	0.671 (0.004)	0.671 (0.004)	0.689 (0.005)	0.689 (0.005)	0.689 (0.005)
0.5	clust	0.810 (0.037)	0.818 (0.028)	0.911 (0.004)	0.536 (0.135)	0.519 (0.149)	0.861 (0.004)	0.638 (0.069)	0.387 (0.267)	0.816 (0.006)
0.5	strict-lassclust	0.790 (0.047)	0.736 (0.090)	0.907 (0.003)	0.518 (0.147)	0.590 (0.122)	0.861 (0.005)	0.630 (0.074)	0.442 (0.222)	0.818 (0.008)
0.5	non-strict-lassclust	0.822 (0.031)	0.832 (0.028)	0.907 (0.003)	0.530 (0.135)	0.591 (0.111)	0.861 (0.005)	0.636 (0.069)	0.660 (0.053)	0.818 (0.008)
0.5	lassojoint	0.827 (0.001)	0.828 (0.002)	0.828 (0.002)	0.429 (0.003)	0.432 (0.004)	0.432 (0.004)	0.587 (0.001)	0.588 (0.001)	0.588 (0.001)
0.8	naive	0.863 (0.000)	0.863 (0.000)	0.863 (0.000)	0.801 (0.000)	0.801 (0.000)	0.801 (0.000)	0.778 (0.000)	0.778 (0.000)	0.778 (0.000)
0.8	clust	0.722 (0.000)	0.739 (0.000)	0.905 (0.000)	0.630 (0.000)	0.666 (0.000)	0.849 (0.000)	0.210 (0.000)	0.413 (0.000)	0.810 (0.000)
0.8	strict-lassclust	0.568 (0.000)	0.772 (0.000)	0.902 (0.000)	0.607 (0.000)	0.709 (0.000)	0.841 (0.000)	0.756 (0.000)	0.685 (0.000)	0.805 (0.000)
0.8	non-strict-lassclust	0.828 (0.000)	0.803 (0.000)	0.902 (0.000)	0.746 (0.000)	0.537 (0.000)	0.842 (0.000)	0.724 (0.000)	0.621 (0.000)	0.805 (0.000)
0.8	lassojoint	0.816 (0.000)	0.816 (0.000)	0.816 (0.000)	0.395 (0.000)	0.395 (0.000)	0.395 (0.000)	0.565 (0.000)	0.565 (0.000)	0.565 (0.000)

Table 14. Summary Statistics for df = wdbc; nonSCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.643 (0.028)	0.643 (0.028)	0.643 (0.028)	0.515 (0.038)	0.515 (0.038)	0.515 (0.038)	0.603 (0.033)	0.603 (0.033)	0.603 (0.033)
0.3	clust	0.889 (0.023)	0.895 (0.025)	0.898 (0.026)	0.879 (0.022)	0.883 (0.023)	0.885 (0.025)	0.822 (0.034)	0.831 (0.036)	0.832 (0.040)
0.3	strict-lassclust	0.946 (0.033)	0.946 (0.024)	0.956 (0.015)	0.839 (0.066)	0.834 (0.075)	0.891 (0.037)	0.728 (0.161)	0.711 (0.189)	0.840 (0.086)
0.3	non-strict-lassclust	0.956 (0.014)	0.955 (0.015)	0.958 (0.014)	0.889 (0.023)	0.888 (0.046)	0.898 (0.024)	0.845 (0.033)	0.850 (0.046)	0.858 (0.033)
0.3	lassojoint	0.963 (0.014)	0.963 (0.013)	0.964 (0.013)	0.395 (0.033)	0.395 (0.034)	0.398 (0.035)	0.552 (0.031)	0.552 (0.031)	0.553 (0.031)
0.5	naive	0.811 (0.004)	0.811 (0.004)	0.811 (0.004)	0.682 (0.004)	0.682 (0.004)	0.682 (0.004)	0.675 (0.001)	0.675 (0.001)	0.675 (0.001)
0.5	clust	0.915 (0.000)	0.923 (0.002)	0.956 (0.004)	0.918 (0.003)	0.906 (0.001)	0.928 (0.005)	0.868 (0.003)	0.849 (0.000)	0.892 (0.008)
0.5	strict-lassclust	0.971 (0.001)	0.961 (0.001)	0.979 (0.001)	0.828 (0.006)	0.726 (0.017)	0.941 (0.003)	0.673 (0.015)	0.304 (0.054)	0.911 (0.004)
0.5	non-strict-lassclust	0.977 (0.001)	0.978 (0.001)	0.978 (0.001)	0.918 (0.000)	0.906 (0.001)	0.941 (0.003)	0.877 (0.001)	0.875 (0.003)	0.911 (0.004)
0.5	lassojoint	0.983 (0.001)	0.983 (0.001)	0.983 (0.001)	0.557 (0.021)	0.557 (0.021)	0.557 (0.021)	0.598 (0.008)	0.598 (0.008)	0.598 (0.008)
0.8	naive	0.919 (0.000)	0.919 (0.000)	0.919 (0.000)	0.906 (0.000)	0.906 (0.000)	0.906 (0.000)	0.881 (0.000)	0.881 (0.000)	0.881 (0.000)
0.8	clust	0.919 (0.000)	0.923 (0.000)	0.960 (0.000)	0.906 (0.000)	0.912 (0.000)	0.929 (0.000)	0.855 (0.000)	0.862 (0.000)	0.895 (0.000)
0.8	strict-lassclust	0.982 (0.000)	0.982 (0.000)	0.987 (0.000)	0.782 (0.000)	0.835 (0.000)	0.935 (0.000)	0.565 (0.000)	0.702 (0.000)	0.906 (0.000)
0.8	non-strict-lassclust	0.985 (0.000)	0.984 (0.000)	0.987 (0.000)	0.941 (0.000)	0.935 (0.000)	0.935 (0.000)	0.921 (0.000)	0.913 (0.000)	0.906 (0.000)
0.8	lassojoint	0.995 (0.000)	0.995 (0.000)	0.995 (0.000)	0.382 (0.000)	0.371 (0.000)	0.382 (0.000)	0.537 (0.000)	0.533 (0.000)	0.537 (0.000)

Table 15. Summary Statistics for df = wine-quality; nonSCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.796 (0.034)	0.796 (0.034)	0.796 (0.034)	0.380 (0.018)	0.380 (0.018)	0.380 (0.018)	0.296 (0.022)	0.296 (0.022)	0.296 (0.022)
0.3	clust	0.649 (0.034)	0.667 (0.034)	0.682 (0.033)	0.685 (0.018)	0.686 (0.019)	0.673 (0.025)	0.325 (0.029)	0.341 (0.029)	0.354 (0.030)
0.3	strict-lassclust	0.646 (0.034)	0.664 (0.034)	0.682 (0.033)	0.685 (0.018)	0.686 (0.020)	0.673 (0.025)	0.325 (0.029)	0.341 (0.028)	0.354 (0.029)
0.3	non-strict-lassclust	0.647 (0.034)	0.665 (0.034)	0.682 (0.033)	0.686 (0.018)	0.686 (0.018)	0.673 (0.024)	0.327 (0.029)	0.341 (0.029)	0.353 (0.029)
0.3	lassojoint	0.813 (0.026)	0.813 (0.027)	0.813 (0.027)	0.211 (0.070)	0.207 (0.069)	0.211 (0.071)	0.254 (0.027)	0.254 (0.027)	0.255 (0.027)
0.5	naive	0.820 (0.017)	0.820 (0.017)	0.820 (0.017)	0.566 (0.005)	0.566 (0.005)	0.566 (0.005)	0.362 (0.009)	0.362 (0.009)	0.362 (0.009)
0.5	clust	0.656 (0.007)	0.714 (0.010)	0.720 (0.010)	0.700 (0.010)	0.682 (0.018)	0.675 (0.019)	0.323 (0.011)	0.370 (0.011)	0.373 (0.016)
0.5	strict-lassclust	0.652 (0.008)	0.726 (0.011)	0.721 (0.009)	0.701 (0.005)	0.737 (0.015)	0.676 (0.012)	0.319 (0.012)	0.398 (0.008)	0.375 (0.013)
0.5	non-strict-lassclust	0.656 (0.006)	0.712 (0.012)	0.721 (0.009)	0.701 (0.009)	0.680 (0.026)	0.678 (0.015)	0.328 (0.005)	0.380 (0.012)	0.376 (0.014)
0.5	lassojoint	0.814 (0.017)	0.814 (0.017)	0.814 (0.018)	0.564 (0.024)	0.555 (0.073)	0.551 (0.072)	0.358 (0.012)	0.356 (0.021)	0.355 (0.020)
0.8	naive	0.854 (0.001)	0.854 (0.001)	0.854 (0.001)	0.754 (0.001)	0.754 (0.001)	0.754 (0.001)	0.483 (0.003)	0.483 (0.003)	0.483 (0.003)
0.8	clust	0.802 (0.007)	0.809 (0.005)	0.817 (0.005)	0.779 (0.002)	0.760 (0.001)	0.666 (0.000)	0.453 (0.005)	0.444 (0.002)	0.407 (0.001)
0.8	strict-lassclust	0.801 (0.008)	0.808 (0.005)	0.818 (0.006)	0.776 (0.003)	0.763 (0.004)	0.650 (0.004)	0.445 (0.004)	0.454 (0.001)	0.397 (0.004)
0.8	non-strict-lassclust	0.801 (0.007)	0.809 (0.005)	0.817 (0.006)	0.772 (0.000)	0.762 (0.001)	0.664 (0.001)	0.440 (0.004)	0.447 (0.002)	0.406 (0.002)
0.8	lassojoint	0.849 (0.001)	0.850 (0.002)	0.850 (0.002)	0.829 (0.002)	0.831 (0.002)	0.831 (0.002)	0.512 (0.001)	0.515 (0.000)	0.515 (0.000)

1.2 nonSCAR scheme - with SMOTE algorithm; pecking_part = 1

Table 16. Summary Statistics for df = adult; nonSCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.769 (0.005)	0.762 (0.006)	0.454 (0.005)	0.462 (0.005)	0.462 (0.006)	0.465 (0.006)
0.3	clust	0.763 (0.006)	0.779 (0.005)	0.621 (0.006)	0.588 (0.005)	0.513 (0.006)	0.510 (0.006)
0.3	strict-lassclust	0.763 (0.006)	0.779 (0.005)	0.621 (0.006)	0.588 (0.005)	0.513 (0.006)	0.510 (0.006)
0.3	non-strict-lassclust	0.763 (0.006)	0.779 (0.005)	0.621 (0.006)	0.588 (0.005)	0.513 (0.006)	0.510 (0.006)
0.3	lassojoint	0.751 (0.014)	0.772 (0.005)	0.363 (0.050)	0.456 (0.005)	0.427 (0.018)	0.462 (0.006)
0.3	spy	0.695 (0.006)		0.375 (0.006)		0.433 (0.006)	
0.5	naive	0.771 (0.003)	0.771 (0.003)	0.578 (0.003)	0.578 (0.003)	0.504 (0.006)	0.504 (0.006)
0.5	clust	0.766 (0.004)	0.768 (0.004)	0.595 (0.001)	0.657 (0.002)	0.511 (0.006)	0.528 (0.007)
0.5	strict-lassclust	0.766 (0.004)	0.768 (0.004)	0.595 (0.001)	0.657 (0.002)	0.511 (0.006)	0.528 (0.007)
0.5	non-strict-lassclust	0.766 (0.004)	0.768 (0.004)	0.595 (0.001)	0.657 (0.002)	0.511 (0.006)	0.528 (0.007)
0.5	lassojoint	0.766 (0.003)	0.766 (0.003)	0.586 (0.024)	0.586 (0.024)	0.501 (0.009)	0.501 (0.009)
0.5	spy	0.695 (0.004)		0.485 (0.015)		0.474 (0.008)	
0.8	naive	0.779 (0.000)	0.779 (0.000)	0.714 (0.002)	0.714 (0.002)	0.543 (0.002)	0.543 (0.002)
0.8	clust	0.770 (0.001)	0.777 (0.000)	0.455 (0.004)	0.729 (0.002)	0.469 (0.003)	0.536 (0.002)
0.8	strict-lassclust	0.770 (0.001)	0.777 (0.000)	0.454 (0.004)	0.729 (0.002)	0.469 (0.003)	0.536 (0.002)
0.8	non-strict-lassclust	0.770 (0.001)	0.777 (0.000)	0.454 (0.004)	0.729 (0.002)	0.469 (0.003)	0.536 (0.002)
0.8	lassojoint	0.775 (0.001)	0.775 (0.001)	0.735 (0.002)	0.735 (0.002)	0.531 (0.002)	0.531 (0.002)
0.8	spy	0.676 (0.006)		0.697 (0.004)		0.231 (0.018)	

Table 17. Summary Statistics for df = artif; nonSCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.764 (0.017)	0.750 (0.017)	0.585 (0.018)	0.617 (0.018)	0.694 (0.016)	0.700 (0.017)
0.3	clust	0.552 (0.039)	0.747 (0.031)	0.518 (0.036)	0.683 (0.023)	0.591 (0.037)	0.703 (0.038)
0.3	strict-lassclust	0.552 (0.039)	0.759 (0.028)	0.520 (0.034)	0.680 (0.027)	0.580 (0.087)	0.699 (0.045)
0.3	non-strict-lassclust	0.553 (0.040)	0.759 (0.028)	0.522 (0.036)	0.683 (0.023)	0.589 (0.040)	0.702 (0.036)
0.3	lassojoint	0.725 (0.021)	0.725 (0.025)	0.550 (0.050)	0.546 (0.045)	0.677 (0.023)	0.674 (0.017)
0.3	spy	0.591 (0.062)		0.492 (0.017)		0.658 (0.015)	
0.5	naive	0.850 (0.011)	0.848 (0.010)	0.674 (0.004)	0.698 (0.003)	0.741 (0.006)	0.749 (0.005)
0.5	clust	0.564 (0.012)	0.782 (0.017)	0.526 (0.003)	0.686 (0.024)	0.633 (0.008)	0.590 (0.065)
0.5	strict-lassclust	0.573 (0.013)	0.779 (0.017)	0.504 (0.014)	0.710 (0.012)	0.648 (0.009)	0.676 (0.023)
0.5	non-strict-lassclust	0.566 (0.008)	0.782 (0.016)	0.524 (0.005)	0.679 (0.026)	0.607 (0.021)	0.563 (0.076)
0.5	lassojoint	0.843 (0.015)	0.835 (0.024)	0.676 (0.006)	0.697 (0.017)	0.739 (0.006)	0.749 (0.003)
0.5	spy	0.698 (0.051)		0.492 (0.017)		0.657 (0.015)	
0.8	naive	0.942 (0.000)	0.942 (0.000)	0.820 (0.000)	0.820 (0.000)	0.841 (0.000)	0.841 (0.000)
0.8	clust	0.641 (0.000)	0.901 (0.000)	0.525 (0.000)	0.768 (0.000)	0.675 (0.000)	0.717 (0.000)
0.8	strict-lassclust	0.653 (0.000)	0.899 (0.000)	0.552 (0.000)	0.803 (0.000)	0.679 (0.000)	0.794 (0.000)
0.8	non-strict-lassclust	0.637 (0.000)	0.903 (0.000)	0.513 (0.000)	0.768 (0.000)	0.670 (0.000)	0.709 (0.000)
0.8	lassojoint	0.847 (0.000)	0.847 (0.000)	0.513 (0.000)	0.513 (0.000)	0.671 (0.000)	0.671 (0.000)
0.8	spy	0.834 (0.016)		0.737 (0.028)		0.751 (0.022)	

Table 18. Summary Statistics for df = bank-marketing; nonSCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.560 (0.028)	0.561 (0.027)	0.300 (0.010)	0.295 (0.011)	0.160 (0.012)	0.127 (0.013)
0.3	clust	0.702 (0.020)	0.704 (0.023)	0.280 (0.011)	0.293 (0.012)	0.088 (0.011)	0.082 (0.011)
0.3	strict-lassclust	0.710 (0.021)	0.709 (0.021)	0.282 (0.011)	0.293 (0.012)	0.084 (0.011)	0.083 (0.011)
0.3	non-strict-lassclust	0.710 (0.021)	0.709 (0.021)	0.282 (0.011)	0.293 (0.012)	0.084 (0.011)	0.083 (0.011)
0.3	lassojoint	0.604 (0.041)	0.591 (0.039)	0.195 (0.089)	0.159 (0.075)	0.182 (0.036)	0.187 (0.038)
0.3	spy	0.705 (0.020)		0.281 (0.011)		0.083 (0.011)	
0.5	naive	0.561 (0.014)	0.561 (0.014)	0.380 (0.005)	0.380 (0.005)	0.126 (0.008)	0.126 (0.008)
0.5	clust	0.433 (0.196)	0.633 (0.104)	0.119 (0.018)	0.526 (0.019)	0.106 (0.018)	0.089 (0.006)
0.5	strict-lassclust	0.776 (0.018)	0.698 (0.018)	0.290 (0.005)	0.562 (0.075)	0.073 (0.006)	0.079 (0.010)
0.5	non-strict-lassclust	0.776 (0.018)	0.698 (0.020)	0.290 (0.005)	0.341 (0.073)	0.073 (0.006)	0.077 (0.010)
0.5	lassojoint	0.612 (0.015)	0.612 (0.015)	0.533 (0.008)	0.533 (0.008)	0.126 (0.013)	0.126 (0.013)
0.5	spy	0.688 (0.015)		0.367 (0.012)		0.081 (0.010)	
0.8	naive	0.648 (0.013)	0.685 (0.012)	0.774 (0.005)	0.738 (0.003)	0.261 (0.009)	0.316 (0.000)
0.8	clust	0.461 (0.009)	0.609 (0.001)	0.107 (0.001)	0.729 (0.002)	0.168 (0.007)	0.159 (0.003)
0.8	strict-lassclust	0.672 (0.003)	0.599 (0.003)	0.300 (0.005)	0.731 (0.002)	0.094 (0.003)	0.156 (0.002)
0.8	non-strict-lassclust	0.672 (0.003)	0.599 (0.003)	0.300 (0.005)	0.731 (0.002)	0.094 (0.003)	0.156 (0.002)
0.8	lassojoint	0.547 (0.013)	0.553 (0.014)	0.773 (0.004)	0.768 (0.003)	0.200 (0.003)	0.246 (0.001)
0.8	spy	0.745 (0.032)		0.702 (0.039)		0.341 (0.039)	

Table 19. Summary Statistics for df = banknote; nonSCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.810 (0.024)	0.755 (0.028)	0.470 (0.022)	0.583 (0.024)	0.620 (0.021)	0.635 (0.025)
0.3	clust	0.728 (0.063)	0.596 (0.033)	0.324 (0.039)	0.463 (0.022)	0.437 (0.033)	0.412 (0.027)
0.3	strict-lassclust	0.722 (0.036)	0.597 (0.033)	0.326 (0.045)	0.462 (0.022)	0.426 (0.026)	0.410 (0.027)
0.3	non-strict-lassclust	0.723 (0.037)	0.598 (0.033)	0.328 (0.045)	0.462 (0.022)	0.428 (0.033)	0.411 (0.027)
0.3	lassojoint	0.603 (0.160)	0.560 (0.085)	0.413 (0.094)	0.452 (0.037)	0.469 (0.158)	0.509 (0.035)
0.3	spy	0.639 (0.031)		0.402 (0.033)		0.431 (0.047)	
0.5	naive	0.915 (0.000)	0.918 (0.000)	0.591 (0.009)	0.761 (0.004)	0.679 (0.007)	0.775 (0.004)
0.5	clust	0.803 (0.002)	0.458 (0.037)	0.341 (0.007)	0.503 (0.005)	0.495 (0.010)	0.424 (0.011)
0.5	strict-lassclust	0.801 (0.001)	0.457 (0.037)	0.341 (0.007)	0.503 (0.005)	0.495 (0.010)	0.424 (0.011)
0.5	non-strict-lassclust	0.801 (0.001)	0.457 (0.037)	0.341 (0.007)	0.503 (0.005)	0.495 (0.010)	0.424 (0.011)
0.5	lassojoint	0.926 (0.000)	0.934 (0.000)	0.509 (0.006)	0.675 (0.013)	0.645 (0.006)	0.727 (0.010)
0.5	spy	0.575 (0.050)		0.483 (0.047)		0.389 (0.074)	
0.8	naive	1.000 (0.000)	1.000 (0.000)	0.990 (0.000)	0.990 (0.000)	0.988 (0.000)	0.988 (0.000)
0.8	clust	0.571 (0.000)	0.829 (0.000)	0.506 (0.000)	0.813 (0.000)	0.609 (0.000)	0.707 (0.000)
0.8	strict-lassclust	0.586 (0.000)	0.838 (0.000)	0.482 (0.000)	0.813 (0.000)	0.600 (0.000)	0.707 (0.000)
0.8	non-strict-lassclust	0.586 (0.000)	0.838 (0.000)	0.482 (0.000)	0.813 (0.000)	0.600 (0.000)	0.707 (0.000)
0.8	lassojoint	0.985 (0.000)	0.985 (0.000)	0.839 (0.000)	0.839 (0.000)	0.835 (0.000)	0.835 (0.000)
0.8	spy	0.747 (0.086)		0.646 (0.048)		0.475 (0.077)	

Table 20. Summary Statistics for df = breastc; nonSCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.924 (0.013)	0.924 (0.016)	0.464 (0.032)	0.669 (0.038)	0.565 (0.031)	0.678 (0.036)
0.3	clust	0.957 (0.021)	0.966 (0.013)	0.953 (0.018)	0.959 (0.014)	0.930 (0.029)	0.940 (0.020)
0.3	strict-lassclust	0.994 (0.003)	0.994 (0.003)	0.959 (0.015)	0.962 (0.013)	0.941 (0.022)	0.944 (0.018)
0.3	non-strict-lassclust	0.994 (0.003)	0.994 (0.003)	0.959 (0.015)	0.962 (0.013)	0.941 (0.022)	0.945 (0.018)
0.3	lassojoint	0.987 (0.007)	0.989 (0.007)	0.364 (0.059)	0.379 (0.097)	0.523 (0.039)	0.532 (0.055)
0.3	spy	0.988 (0.007)		0.918 (0.025)		0.895 (0.031)	
0.5	naive	0.938 (0.002)	0.938 (0.002)	0.753 (0.014)	0.821 (0.007)	0.744 (0.014)	0.800 (0.008)
0.5	clust	0.976 (0.007)	0.978 (0.002)	0.970 (0.004)	0.971 (0.001)	0.956 (0.006)	0.958 (0.002)
0.5	strict-lassclust	0.998 (0.000)	0.998 (0.000)	0.975 (0.001)	0.966 (0.001)	0.966 (0.002)	0.952 (0.002)
0.5	non-strict-lassclust	0.998 (0.000)	0.998 (0.000)	0.975 (0.001)	0.966 (0.001)	0.966 (0.002)	0.952 (0.002)
0.5	lassojoint	0.998 (0.001)	0.998 (0.001)	0.359 (0.006)	0.359 (0.006)	0.528 (0.006)	0.528 (0.006)
0.5	spy	0.989 (0.006)		0.920 (0.022)		0.897 (0.027)	
0.8	naive	0.990 (0.001)	0.990 (0.001)	0.935 (0.003)	0.935 (0.003)	0.910 (0.004)	0.910 (0.004)
0.8	clust	0.969 (0.002)	0.967 (0.002)	0.969 (0.007)	0.971 (0.001)	0.952 (0.011)	0.955 (0.002)
0.8	strict-lassclust	0.996 (0.000)	0.996 (0.000)	0.960 (0.006)	0.971 (0.001)	0.937 (0.009)	0.955 (0.001)
0.8	non-strict-lassclust	0.996 (0.000)	0.996 (0.000)	0.960 (0.006)	0.971 (0.001)	0.937 (0.009)	0.955 (0.001)
0.8	lassojoint	0.993 (0.000)	0.993 (0.000)	0.854 (0.034)	0.854 (0.034)	0.819 (0.041)	0.819 (0.041)
0.8	spy	0.989 (0.007)		0.903 (0.020)		0.878 (0.025)	

Table 21. Summary Statistics for df = credit-a; nonSCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.714 (0.035)	0.708 (0.035)	0.542 (0.027)	0.563 (0.027)	0.647 (0.028)	0.649 (0.027)
0.3	clust	0.622 (0.027)	0.657 (0.029)	0.644 (0.029)	0.674 (0.030)	0.413 (0.056)	0.509 (0.057)
0.3	strict-lassclust	0.715 (0.040)	0.726 (0.033)	0.643 (0.030)	0.674 (0.030)	0.410 (0.057)	0.509 (0.058)
0.3	non-strict-lassclust	0.715 (0.040)	0.726 (0.033)	0.643 (0.030)	0.674 (0.030)	0.410 (0.057)	0.509 (0.058)
0.3	lassojoint	0.679 (0.035)	0.679 (0.037)	0.529 (0.045)	0.488 (0.057)	0.639 (0.028)	0.632 (0.031)
0.3	spy	0.735 (0.044)		0.682 (0.036)		0.641 (0.052)	
0.5	naive	0.721 (0.008)	0.719 (0.008)	0.574 (0.006)	0.610 (0.008)	0.639 (0.006)	0.652 (0.006)
0.5	clust	0.589 (0.007)	0.601 (0.007)	0.616 (0.004)	0.623 (0.006)	0.353 (0.017)	0.379 (0.022)
0.5	strict-lassclust	0.728 (0.008)	0.748 (0.015)	0.622 (0.003)	0.622 (0.005)	0.366 (0.013)	0.379 (0.021)
0.5	non-strict-lassclust	0.728 (0.008)	0.748 (0.015)	0.622 (0.003)	0.622 (0.005)	0.366 (0.013)	0.379 (0.021)
0.5	lassojoint	0.705 (0.010)	0.669 (0.011)	0.609 (0.026)	0.475 (0.019)	0.653 (0.005)	0.622 (0.009)
0.5	spy	0.737 (0.042)		0.686 (0.032)		0.651 (0.043)	
0.8	naive	0.827 (0.000)	0.827 (0.000)	0.687 (0.000)	0.687 (0.000)	0.687 (0.000)	0.687 (0.000)
0.8	clust	0.649 (0.000)	0.649 (0.000)	0.703 (0.000)	0.703 (0.000)	0.463 (0.000)	0.463 (0.000)
0.8	strict-lassclust	0.794 (0.000)	0.792 (0.000)	0.703 (0.000)	0.697 (0.000)	0.453 (0.000)	0.449 (0.000)
0.8	non-strict-lassclust	0.794 (0.000)	0.792 (0.000)	0.703 (0.000)	0.697 (0.000)	0.453 (0.000)	0.449 (0.000)
0.8	lassojoint	0.824 (0.000)	0.824 (0.000)	0.738 (0.000)	0.738 (0.000)	0.721 (0.000)	0.721 (0.000)
0.8	spy	0.751 (0.048)		0.699 (0.039)		0.618 (0.071)	

Table 22. Summary Statistics for df = credit-g; nonSCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.579 (0.036)	0.587 (0.034)	0.340 (0.023)	0.399 (0.025)	0.420 (0.028)	0.380 (0.032)
0.3	clust	0.567 (0.044)	0.543 (0.038)	0.600 (0.034)	0.625 (0.050)	0.387 (0.039)	0.325 (0.043)
0.3	strict-lassclust	0.572 (0.042)	0.544 (0.038)	0.617 (0.033)	0.629 (0.051)	0.380 (0.040)	0.324 (0.044)
0.3	non-strict-lassclust	0.571 (0.043)	0.542 (0.038)	0.605 (0.051)	0.624 (0.051)	0.383 (0.039)	0.326 (0.043)
0.3	lassojoint	0.602 (0.034)	0.603 (0.034)	0.394 (0.027)	0.394 (0.026)	0.344 (0.033)	0.344 (0.032)
0.3	spy	0.574 (0.050)		0.473 (0.031)		0.243 (0.037)	
0.5	naive	0.569 (0.024)	0.569 (0.024)	0.408 (0.002)	0.408 (0.002)	0.371 (0.024)	0.371 (0.024)
0.5	clust	0.519 (0.021)	0.565 (0.024)	0.492 (0.046)	0.683 (0.014)	0.367 (0.006)	0.283 (0.038)
0.5	strict-lassclust	0.522 (0.019)	0.569 (0.022)	0.498 (0.044)	0.686 (0.013)	0.370 (0.005)	0.281 (0.045)
0.5	non-strict-lassclust	0.522 (0.019)	0.564 (0.033)	0.498 (0.044)	0.679 (0.031)	0.370 (0.005)	0.295 (0.017)
0.5	lassojoint	0.625 (0.013)	0.624 (0.012)	0.437 (0.007)	0.437 (0.007)	0.255 (0.030)	0.255 (0.030)
0.5	spy	0.578 (0.050)		0.465 (0.024)		0.238 (0.036)	
0.8	naive	0.524 (0.009)	0.524 (0.009)	0.579 (0.005)	0.579 (0.005)	0.280 (0.018)	0.280 (0.018)
0.8	clust	0.528 (0.029)	0.595 (0.013)	0.363 (0.013)	0.717 (0.008)	0.423 (0.010)	0.285 (0.024)
0.8	strict-lassclust	0.540 (0.018)	0.598 (0.015)	0.364 (0.008)	0.717 (0.008)	0.420 (0.008)	0.285 (0.024)
0.8	non-strict-lassclust	0.540 (0.018)	0.598 (0.015)	0.364 (0.008)	0.717 (0.008)	0.420 (0.008)	0.285 (0.024)
0.8	lassojoint	0.643 (0.012)	0.642 (0.010)	0.610 (0.000)	0.610 (0.000)	0.033 (0.000)	0.033 (0.000)
0.8	spy	0.573 (0.055)		0.581 (0.086)		0.375 (0.059)	

Table 23. Summary Statistics for df = dhfr; nonSCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.538 (0.048)	0.548 (0.048)	0.466 (0.051)	0.472 (0.056)	0.481 (0.059)	0.499 (0.058)
0.3	clust	0.567 (0.120)	0.543 (0.056)	0.357 (0.096)	0.532 (0.070)	0.422 (0.068)	0.384 (0.090)
0.3	strict-lassclust	0.590 (0.079)	0.560 (0.063)	0.446 (0.065)	0.550 (0.058)	0.381 (0.088)	0.324 (0.082)
0.3	non-strict-lassclust	0.582 (0.070)	0.578 (0.060)	0.449 (0.054)	0.564 (0.060)	0.392 (0.079)	0.329 (0.090)
0.3	lassojoint	0.660 (0.064)	0.636 (0.050)	0.504 (0.085)	0.473 (0.049)	0.562 (0.025)	0.559 (0.037)
0.3	spy	0.658 (0.098)		0.656 (0.091)		0.519 (0.125)	
0.5	naive	0.728 (0.000)	0.730 (0.000)	0.629 (0.000)	0.680 (0.000)	0.667 (0.000)	0.699 (0.000)
0.5	clust	0.552 (0.000)	0.596 (0.000)	0.454 (0.000)	0.577 (0.000)	0.505 (0.000)	0.494 (0.000)
0.5	strict-lassclust	0.594 (0.000)	0.532 (0.000)	0.371 (0.000)	0.588 (0.000)	0.396 (0.000)	0.429 (0.000)
0.5	non-strict-lassclust	0.592 (0.000)	0.536 (0.000)	0.351 (0.000)	0.515 (0.000)	0.364 (0.000)	0.390 (0.000)
0.5	lassojoint	0.704 (0.000)	0.710 (0.000)	0.464 (0.000)	0.629 (0.000)	0.612 (0.000)	0.667 (0.000)
0.5	spy	0.680 (0.057)		0.665 (0.039)		0.498 (0.038)	
0.8	naive	0.653 (0.000)	0.653 (0.000)	0.608 (0.000)	0.608 (0.000)	0.578 (0.000)	0.578 (0.000)
0.8	clust	0.512 (0.000)	0.713 (0.000)	0.412 (0.000)	0.691 (0.000)	0.571 (0.000)	0.615 (0.000)
0.8	strict-lassclust	0.610 (0.000)	0.748 (0.000)	0.454 (0.000)	0.639 (0.000)	0.583 (0.000)	0.624 (0.000)
0.8	non-strict-lassclust	0.581 (0.000)	0.761 (0.000)	0.423 (0.000)	0.722 (0.000)	0.582 (0.000)	0.557 (0.000)
0.8	lassojoint	0.845 (0.000)	0.839 (0.000)	0.732 (0.000)	0.691 (0.000)	0.740 (0.000)	0.706 (0.000)
0.8	spy	0.685 (0.084)		0.645 (0.078)		0.534 (0.142)	

Table 24. Summary Statistics for df = diabetes; nonSCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.705 (0.028)	0.701 (0.028)	0.500 (0.029)	0.527 (0.029)	0.566 (0.033)	0.577 (0.033)
0.3	clust	0.690 (0.055)	0.725 (0.051)	0.676 (0.025)	0.694 (0.027)	0.564 (0.047)	0.603 (0.040)
0.3	strict-lassclust	0.722 (0.040)	0.747 (0.035)	0.677 (0.026)	0.694 (0.027)	0.562 (0.051)	0.603 (0.040)
0.3	non-strict-lassclust	0.723 (0.041)	0.747 (0.035)	0.675 (0.028)	0.695 (0.027)	0.567 (0.049)	0.603 (0.040)
0.3	lassojoint	0.717 (0.027)	0.722 (0.029)	0.379 (0.032)	0.552 (0.070)	0.526 (0.031)	0.592 (0.039)
0.3	spy	0.722 (0.029)		0.637 (0.028)		0.615 (0.034)	
0.5	naive	0.700 (0.012)	0.700 (0.012)	0.595 (0.013)	0.599 (0.018)	0.604 (0.008)	0.605 (0.007)
0.5	clust	0.628 (0.013)	0.644 (0.019)	0.656 (0.007)	0.653 (0.008)	0.521 (0.013)	0.512 (0.028)
0.5	strict-lassclust	0.692 (0.012)	0.697 (0.008)	0.653 (0.007)	0.656 (0.007)	0.522 (0.036)	0.514 (0.029)
0.5	non-strict-lassclust	0.690 (0.011)	0.698 (0.008)	0.655 (0.009)	0.658 (0.006)	0.521 (0.034)	0.519 (0.023)
0.5	lassojoint	0.693 (0.009)	0.693 (0.012)	0.454 (0.120)	0.451 (0.123)	0.548 (0.053)	0.545 (0.052)
0.5	spy	0.719 (0.031)		0.644 (0.025)		0.615 (0.032)	
0.8	naive	0.758 (0.010)	0.758 (0.010)	0.697 (0.015)	0.697 (0.015)	0.627 (0.020)	0.627 (0.020)
0.8	clust	0.695 (0.020)	0.716 (0.021)	0.613 (0.043)	0.671 (0.020)	0.575 (0.004)	0.509 (0.045)
0.8	strict-lassclust	0.693 (0.020)	0.716 (0.022)	0.610 (0.041)	0.665 (0.022)	0.566 (0.012)	0.499 (0.051)
0.8	non-strict-lassclust	0.696 (0.020)	0.716 (0.021)	0.607 (0.040)	0.664 (0.021)	0.566 (0.011)	0.497 (0.048)
0.8	lassojoint	0.759 (0.010)	0.760 (0.008)	0.696 (0.005)	0.694 (0.010)	0.620 (0.010)	0.617 (0.015)
0.8	spy	0.712 (0.030)		0.662 (0.027)		0.614 (0.033)	

Table 25. Summary Statistics for df = heart-c; nonSCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.584 (0.056)	0.562 (0.048)	0.494 (0.042)	0.517 (0.044)	0.622 (0.040)	0.582 (0.045)
0.3	clust	0.693 (0.047)	0.664 (0.045)	0.627 (0.045)	0.639 (0.040)	0.666 (0.045)	0.634 (0.047)
0.3	strict-lassclust	0.692 (0.048)	0.663 (0.045)	0.624 (0.047)	0.638 (0.040)	0.665 (0.044)	0.632 (0.047)
0.3	non-strict-lassclust	0.692 (0.048)	0.663 (0.045)	0.624 (0.047)	0.638 (0.040)	0.665 (0.044)	0.632 (0.047)
0.3	lassojoint	0.565 (0.062)	0.592 (0.075)	0.467 (0.046)	0.486 (0.061)	0.538 (0.065)	0.548 (0.080)
0.3	spy	0.688 (0.045)		0.628 (0.046)		0.655 (0.050)	
0.5	naive	0.570 (0.000)	0.551 (0.000)	0.578 (0.000)	0.578 (0.000)	0.678 (0.000)	0.661 (0.000)
0.5	clust	0.640 (0.000)	0.642 (0.000)	0.600 (0.000)	0.633 (0.000)	0.700 (0.000)	0.667 (0.000)
0.5	strict-lassclust	0.635 (0.000)	0.636 (0.000)	0.600 (0.000)	0.622 (0.000)	0.700 (0.000)	0.653 (0.000)
0.5	non-strict-lassclust	0.635 (0.000)	0.636 (0.000)	0.600 (0.000)	0.622 (0.000)	0.700 (0.000)	0.653 (0.000)
0.5	lassojoint	0.594 (0.000)	0.642 (0.000)	0.589 (0.000)	0.622 (0.000)	0.694 (0.000)	0.712 (0.000)
0.5	spy	0.681 (0.052)		0.636 (0.048)		0.615 (0.137)	
0.8	naive	0.754 (0.000)	0.754 (0.000)	0.667 (0.000)	0.667 (0.000)	0.732 (0.000)	0.732 (0.000)
0.8	clust	0.685 (0.000)	0.733 (0.000)	0.567 (0.000)	0.700 (0.000)	0.715 (0.000)	0.710 (0.000)
0.8	strict-lassclust	0.686 (0.000)	0.729 (0.000)	0.567 (0.000)	0.678 (0.000)	0.715 (0.000)	0.681 (0.000)
0.8	non-strict-lassclust	0.686 (0.000)	0.729 (0.000)	0.567 (0.000)	0.678 (0.000)	0.715 (0.000)	0.681 (0.000)
0.8	lassojoint	0.759 (0.000)	0.759 (0.000)	0.700 (0.000)	0.700 (0.000)	0.773 (0.000)	0.773 (0.000)
0.8	spy	0.684 (0.051)		0.618 (0.047)		0.480 (0.145)	

Table 26. Summary Statistics for df = qsar-biodeg; nonSCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.633 (0.026)	0.628 (0.027)	0.484 (0.027)	0.502 (0.028)	0.547 (0.030)	0.546 (0.031)
0.3	clust	0.761 (0.034)	0.775 (0.058)	0.575 (0.050)	0.685 (0.066)	0.582 (0.055)	0.599 (0.151)
0.3	strict-lassclust	0.763 (0.035)	0.779 (0.059)	0.623 (0.091)	0.666 (0.115)	0.515 (0.153)	0.571 (0.144)
0.3	non-strict-lassclust	0.739 (0.052)	0.774 (0.058)	0.490 (0.093)	0.651 (0.098)	0.548 (0.078)	0.592 (0.146)
0.3	lassojoint	0.662 (0.033)	0.643 (0.030)	0.437 (0.031)	0.402 (0.029)	0.534 (0.028)	0.524 (0.029)
0.3	spy	0.707 (0.099)		0.525 (0.115)		0.455 (0.173)	
0.5	naive	0.712 (0.001)	0.710 (0.002)	0.561 (0.013)	0.577 (0.027)	0.560 (0.002)	0.562 (0.002)
0.5	clust	0.763 (0.009)	0.822 (0.007)	0.548 (0.012)	0.761 (0.014)	0.572 (0.013)	0.677 (0.025)
0.5	strict-lassclust	0.763 (0.010)	0.824 (0.006)	0.536 (0.039)	0.772 (0.009)	0.571 (0.022)	0.668 (0.028)
0.5	non-strict-lassclust	0.760 (0.008)	0.821 (0.006)	0.475 (0.040)	0.752 (0.025)	0.550 (0.008)	0.677 (0.026)
0.5	lassojoint	0.707 (0.005)	0.692 (0.008)	0.463 (0.038)	0.438 (0.074)	0.540 (0.003)	0.531 (0.016)
0.5	spy	0.744 (0.113)		0.617 (0.092)		0.570 (0.132)	
0.8	naive	0.815 (0.007)	0.815 (0.007)	0.734 (0.000)	0.734 (0.000)	0.667 (0.000)	0.667 (0.000)
0.8	clust	0.779 (0.009)	0.865 (0.004)	0.513 (0.007)	0.813 (0.010)	0.566 (0.002)	0.729 (0.009)
0.8	strict-lassclust	0.765 (0.009)	0.868 (0.008)	0.444 (0.019)	0.812 (0.011)	0.538 (0.007)	0.710 (0.018)
0.8	non-strict-lassclust	0.765 (0.010)	0.870 (0.006)	0.426 (0.008)	0.752 (0.068)	0.531 (0.003)	0.693 (0.034)
0.8	lassojoint	0.750 (0.011)	0.744 (0.005)	0.592 (0.030)	0.585 (0.023)	0.574 (0.011)	0.573 (0.011)
0.8	spy	0.849 (0.023)		0.690 (0.027)		0.661 (0.027)	

Table 27. Summary Statistics for df = spambase; nonSCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.685 (0.010)	0.685 (0.009)	0.552 (0.016)	0.637 (0.011)	0.632 (0.013)	0.668 (0.012)
0.3	clust	0.876 (0.066)	0.915 (0.021)	0.814 (0.113)	0.849 (0.009)	0.777 (0.063)	0.780 (0.014)
0.3	strict-lassclust	0.905 (0.024)	0.920 (0.007)	0.802 (0.135)	0.848 (0.009)	0.772 (0.073)	0.779 (0.015)
0.3	non-strict-lassclust	0.904 (0.026)	0.919 (0.008)	0.801 (0.134)	0.847 (0.009)	0.772 (0.073)	0.778 (0.014)
0.3	lassojoint	0.806 (0.016)	0.819 (0.010)	0.433 (0.017)	0.566 (0.036)	0.579 (0.013)	0.636 (0.019)
0.3	spy	0.797 (0.032)		0.618 (0.061)		0.652 (0.039)	
0.5	naive	0.786 (0.004)	0.786 (0.004)	0.671 (0.004)	0.671 (0.004)	0.689 (0.005)	0.689 (0.005)
0.5	clust	0.911 (0.004)	0.920 (0.006)	0.861 (0.004)	0.843 (0.003)	0.816 (0.006)	0.783 (0.006)
0.5	strict-lassclust	0.907 (0.003)	0.920 (0.004)	0.861 (0.005)	0.848 (0.007)	0.818 (0.008)	0.789 (0.014)
0.5	non-strict-lassclust	0.907 (0.003)	0.916 (0.005)	0.861 (0.005)	0.847 (0.005)	0.818 (0.008)	0.788 (0.012)
0.5	lassojoint	0.828 (0.002)	0.827 (0.001)	0.432 (0.004)	0.429 (0.003)	0.588 (0.001)	0.587 (0.001)
0.5	spy	0.819 (0.044)		0.684 (0.082)		0.684 (0.055)	
0.8	naive	0.863 (0.000)	0.863 (0.000)	0.801 (0.000)	0.801 (0.000)	0.778 (0.000)	0.778 (0.000)
0.8	clust	0.905 (0.000)	0.926 (0.000)	0.849 (0.000)	0.855 (0.000)	0.810 (0.000)	0.790 (0.000)
0.8	strict-lassclust	0.902 (0.000)	0.921 (0.000)	0.841 (0.000)	0.857 (0.000)	0.805 (0.000)	0.793 (0.000)
0.8	non-strict-lassclust	0.902 (0.000)	0.921 (0.000)	0.842 (0.000)	0.857 (0.000)	0.805 (0.000)	0.793 (0.000)
0.8	lassojoint	0.816 (0.000)	0.819 (0.000)	0.395 (0.000)	0.395 (0.000)	0.565 (0.000)	0.565 (0.000)
0.8	spy	0.828 (0.018)		0.737 (0.071)		0.727 (0.046)	

Table 28. Summary Statistics for df = wdbc; nonSCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.643 (0.028)	0.645 (0.027)	0.515 (0.038)	0.516 (0.038)	0.603 (0.033)	0.603 (0.033)
0.3	clust	0.898 (0.026)	0.917 (0.022)	0.885 (0.025)	0.902 (0.021)	0.832 (0.040)	0.867 (0.029)
0.3	strict-lassclust	0.956 (0.015)	0.957 (0.013)	0.891 (0.037)	0.904 (0.021)	0.840 (0.086)	0.868 (0.033)
0.3	non-strict-lassclust	0.958 (0.014)	0.957 (0.013)	0.898 (0.024)	0.904 (0.025)	0.858 (0.033)	0.872 (0.030)
0.3	lassojoint	0.964 (0.013)	0.958 (0.016)	0.398 (0.035)	0.385 (0.033)	0.553 (0.031)	0.548 (0.030)
0.3	spy	0.955 (0.012)		0.895 (0.019)		0.866 (0.026)	
0.5	naive	0.811 (0.004)	0.807 (0.004)	0.682 (0.004)	0.676 (0.003)	0.675 (0.001)	0.671 (0.000)
0.5	clust	0.956 (0.004)	0.940 (0.001)	0.928 (0.005)	0.900 (0.001)	0.892 (0.008)	0.847 (0.003)
0.5	strict-lassclust	0.979 (0.001)	0.976 (0.000)	0.941 (0.003)	0.935 (0.001)	0.911 (0.004)	0.903 (0.000)
0.5	non-strict-lassclust	0.978 (0.001)	0.977 (0.000)	0.941 (0.003)	0.935 (0.001)	0.911 (0.004)	0.903 (0.000)
0.5	lassojoint	0.983 (0.001)	0.974 (0.000)	0.557 (0.021)	0.335 (0.002)	0.598 (0.008)	0.498 (0.003)
0.5	spy	0.958 (0.011)		0.896 (0.016)		0.866 (0.022)	
0.8	naive	0.919 (0.000)	0.919 (0.000)	0.906 (0.000)	0.906 (0.000)	0.881 (0.000)	0.881 (0.000)
0.8	clust	0.960 (0.000)	0.960 (0.000)	0.929 (0.000)	0.935 (0.000)	0.895 (0.000)	0.906 (0.000)
0.8	strict-lassclust	0.987 (0.000)	0.987 (0.000)	0.935 (0.000)	0.935 (0.000)	0.906 (0.000)	0.906 (0.000)
0.8	non-strict-lassclust	0.987 (0.000)	0.987 (0.000)	0.935 (0.000)	0.935 (0.000)	0.906 (0.000)	0.906 (0.000)
0.8	lassojoint	0.995 (0.000)	0.997 (0.000)	0.382 (0.000)	0.388 (0.000)	0.537 (0.000)	0.540 (0.000)
0.8	spy	0.959 (0.013)		0.905 (0.019)		0.875 (0.027)	

Table 29. Summary Statistics for df = wine-quality; nonSCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.796 (0.034)	0.782 (0.032)	0.380 (0.018)	0.387 (0.017)	0.296 (0.022)	0.299 (0.022)
0.3	clust	0.682 (0.033)	0.701 (0.033)	0.673 (0.025)	0.651 (0.020)	0.354 (0.030)	0.357 (0.030)
0.3	strict-lassclust	0.682 (0.033)	0.702 (0.033)	0.673 (0.025)	0.651 (0.020)	0.354 (0.029)	0.357 (0.030)
0.3	non-strict-lassclust	0.682 (0.033)	0.702 (0.033)	0.673 (0.024)	0.652 (0.020)	0.353 (0.029)	0.358 (0.030)
0.3	lassojoint	0.813 (0.027)	0.821 (0.025)	0.211 (0.071)	0.397 (0.039)	0.255 (0.027)	0.301 (0.024)
0.3	spy	0.664 (0.035)		0.677 (0.024)		0.301 (0.039)	
0.5	naive	0.820 (0.017)	0.820 (0.017)	0.566 (0.005)	0.566 (0.005)	0.362 (0.009)	0.362 (0.009)
0.5	clust	0.720 (0.010)	0.720 (0.009)	0.675 (0.019)	0.672 (0.015)	0.373 (0.016)	0.369 (0.017)
0.5	strict-lassclust	0.721 (0.009)	0.721 (0.009)	0.676 (0.012)	0.676 (0.012)	0.375 (0.013)	0.376 (0.014)
0.5	non-strict-lassclust	0.721 (0.009)	0.721 (0.009)	0.678 (0.015)	0.677 (0.013)	0.376 (0.014)	0.376 (0.014)
0.5	lassojoint	0.814 (0.018)	0.814 (0.017)	0.551 (0.072)	0.564 (0.022)	0.355 (0.020)	0.358 (0.012)
0.5	spy	0.633 (0.030)		0.686 (0.017)		0.268 (0.035)	
0.8	naive	0.854 (0.001)	0.858 (0.001)	0.754 (0.001)	0.727 (0.001)	0.483 (0.003)	0.462 (0.004)
0.8	clust	0.817 (0.005)	0.819 (0.007)	0.666 (0.000)	0.801 (0.003)	0.407 (0.001)	0.474 (0.004)
0.8	strict-lassclust	0.818 (0.006)	0.818 (0.007)	0.650 (0.004)	0.803 (0.004)	0.397 (0.004)	0.477 (0.005)
0.8	non-strict-lassclust	0.817 (0.006)	0.818 (0.007)	0.664 (0.001)	0.803 (0.004)	0.406 (0.002)	0.477 (0.005)
0.8	lassojoint	0.850 (0.002)	0.854 (0.004)	0.831 (0.002)	0.787 (0.001)	0.515 (0.000)	0.514 (0.002)
0.8	spy	0.646 (0.032)		0.713 (0.019)		0.292 (0.043)	

1.3 Graphs

In this section, we present all boxplots for classification metrics for our methods. In addition, we present boxplots of the execution times of the algorithms used.

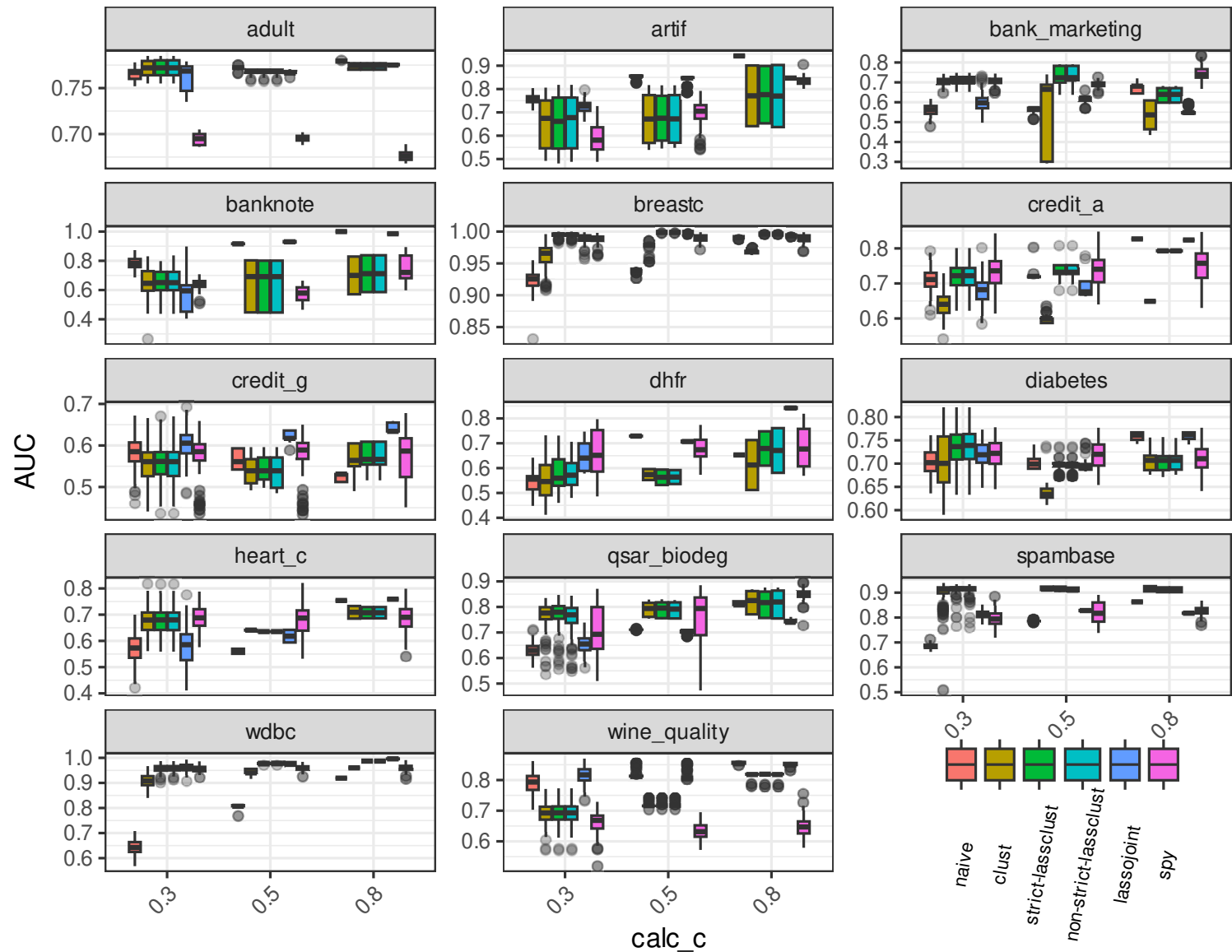


Fig. 1. AUC; SMOTE scheme; nonSCAR

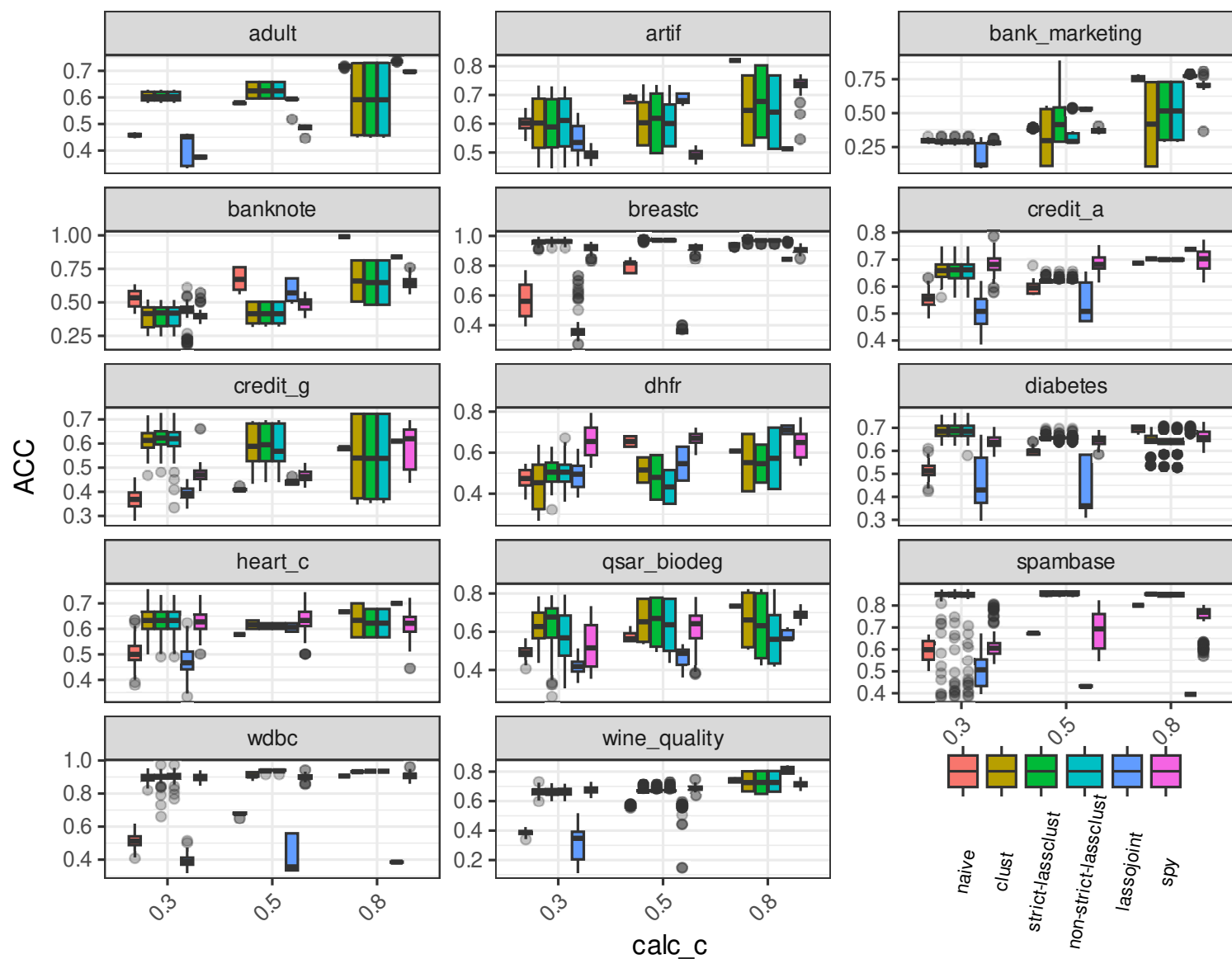


Fig. 2. Accuracy; SMOTE scheme; nonSCAR

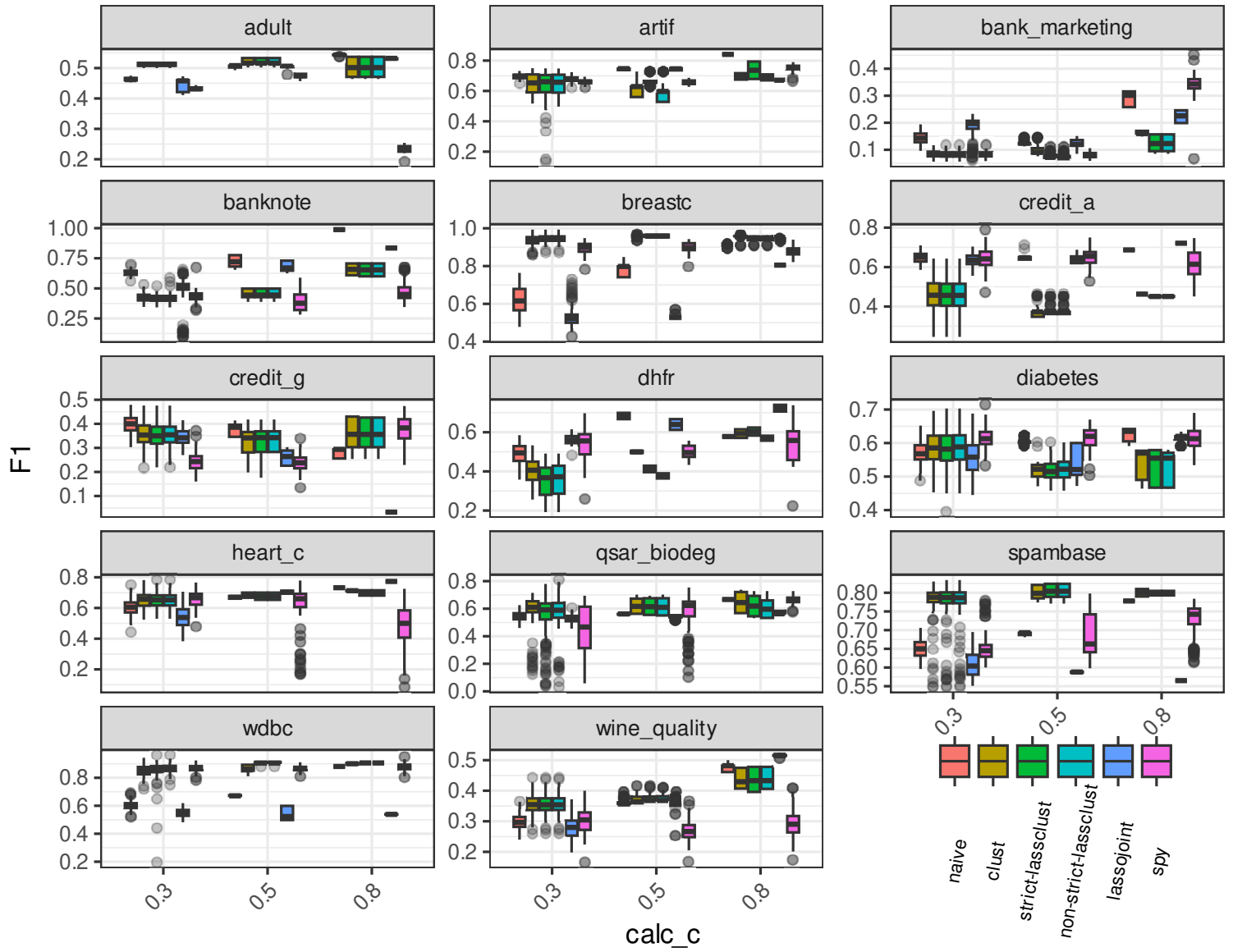


Fig. 3. F1 score; SMOTE scheme; nonSCAR

2 SCAR Scheme

2.1 Tables for all classification metrics for 14 datasets.

q	naive acc	clust acc	strict lassclust acc	non strict lassclust acc	lassojoint acc	naive f1	clust f1	strict lassclust f1	non strict lassclust f1	lassojoint f1	naive auc	clust auc	strict lassclust auc	non strict lassclust auc	lassojoint auc
0.25	0.58 (0.23)	0.7 (0.13)	0.7 (0.13)	0.7 (0.13)	0.75 (0.21)	0.61 (0.18)	0.58 (0.2)	0.57 (0.22)	0.58 (0.2)	0.68 (0.22)	0.88 (0.1)	0.75 (0.12)	0.77 (0.12)	0.76 (0.13)	0.88 (0.11)
0.5	0.58 (0.23)	0.71 (0.13)	0.7 (0.12)	0.7 (0.13)	0.77 (0.2)	0.61 (0.18)	0.58 (0.2)	0.55 (0.21)	0.58 (0.2)	0.69 (0.22)	0.88 (0.1)	0.76 (0.12)	0.76 (0.12)	0.77 (0.13)	0.88 (0.11)
1	0.58 (0.23)	0.6 (0.21)	0.6 (0.21)	0.59 (0.21)	0.77 (0.2)	0.61 (0.18)	0.59 (0.21)	0.58 (0.21)	0.58 (0.21)	0.69 (0.22)	0.88 (0.1)	0.76 (0.12)	0.77 (0.13)	0.76 (0.13)	0.88 (0.11)

Table 30. Summary of Mean and Standard Deviation for Accuracy, F1, and AUC Metrics by q (Highest Mean in Bold); SCAR scheme

Table 31. Summary Statistics for df = adult; SCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.797 (0.003)	0.797 (0.003)	0.797 (0.003)	0.241 (0.005)	0.241 (0.005)	0.241 (0.005)	0.388 (0.006)	0.388 (0.006)	0.388 (0.006)
0.3	clust	0.767 (0.004)	0.766 (0.004)	0.763 (0.005)	0.582 (0.007)	0.577 (0.006)	0.406 (0.011)	0.502 (0.007)	0.501 (0.007)	0.442 (0.008)
0.3	strict-lassclust	0.768 (0.004)	0.766 (0.004)	0.763 (0.005)	0.582 (0.007)	0.577 (0.006)	0.405 (0.011)	0.503 (0.007)	0.500 (0.007)	0.442 (0.008)
0.3	non-strict-lassclust	0.768 (0.004)	0.766 (0.004)	0.763 (0.005)	0.582 (0.007)	0.577 (0.006)	0.405 (0.011)	0.503 (0.007)	0.500 (0.007)	0.442 (0.008)
0.3	lassojoint	0.795 (0.003)	0.795 (0.003)	0.795 (0.003)	0.792 (0.003)	0.792 (0.003)	0.792 (0.003)	0.480 (0.033)	0.480 (0.033)	0.480 (0.033)
0.5	naive	0.804 (0.001)	0.804 (0.001)	0.804 (0.001)	0.335 (0.006)	0.335 (0.006)	0.335 (0.006)	0.419 (0.006)	0.419 (0.006)	0.419 (0.006)
0.5	clust	0.771 (0.003)	0.767 (0.004)	0.761 (0.004)	0.568 (0.004)	0.554 (0.002)	0.304 (0.002)	0.497 (0.006)	0.492 (0.004)	0.407 (0.004)
0.5	strict-lassclust	0.771 (0.003)	0.768 (0.004)	0.761 (0.004)	0.567 (0.004)	0.554 (0.002)	0.302 (0.002)	0.497 (0.006)	0.492 (0.005)	0.407 (0.005)
0.5	non-strict-lassclust	0.771 (0.003)	0.768 (0.004)	0.761 (0.004)	0.567 (0.004)	0.554 (0.002)	0.302 (0.002)	0.497 (0.006)	0.492 (0.005)	0.407 (0.005)
0.5	lassojoint	0.804 (0.002)	0.804 (0.002)	0.804 (0.002)	0.794 (0.005)	0.794 (0.005)	0.794 (0.005)	0.504 (0.016)	0.504 (0.016)	0.504 (0.016)
0.8	naive	0.802 (0.001)	0.802 (0.001)	0.802 (0.001)	0.783 (0.001)	0.783 (0.001)	0.783 (0.001)	0.544 (0.001)	0.544 (0.001)	0.544 (0.001)
0.8	clust	0.769 (0.003)	0.762 (0.001)	0.755 (0.001)	0.605 (0.009)	0.542 (0.002)	0.241 (0.003)	0.499 (0.007)	0.480 (0.003)	0.383 (0.003)
0.8	strict-lassclust	0.769 (0.003)	0.762 (0.001)	0.755 (0.001)	0.604 (0.010)	0.541 (0.002)	0.241 (0.002)	0.499 (0.007)	0.480 (0.002)	0.383 (0.003)
0.8	non-strict-lassclust	0.769 (0.003)	0.762 (0.001)	0.755 (0.001)	0.604 (0.010)	0.541 (0.002)	0.241 (0.002)	0.499 (0.007)	0.480 (0.002)	0.383 (0.003)
0.8	lassojoint	0.801 (0.002)	0.801 (0.002)	0.801 (0.002)	0.796 (0.002)	0.796 (0.002)	0.796 (0.002)	0.470 (0.003)	0.470 (0.003)	0.470 (0.003)

Table 32. Summary Statistics for df = artif; SCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.968 (0.010)	0.968 (0.010)	0.968 (0.010)	0.520 (0.019)	0.520 (0.019)	0.520 (0.019)	0.670 (0.016)	0.670 (0.016)	0.670 (0.016)
0.3	clust	0.604 (0.064)	0.610 (0.063)	0.599 (0.063)	0.571 (0.050)	0.573 (0.052)	0.543 (0.037)	0.559 (0.049)	0.563 (0.054)	0.631 (0.035)
0.3	strict-lassclust	0.602 (0.064)	0.607 (0.062)	0.597 (0.065)	0.559 (0.050)	0.557 (0.049)	0.537 (0.043)	0.557 (0.116)	0.561 (0.105)	0.617 (0.057)
0.3	non-strict-lassclust	0.609 (0.063)	0.613 (0.064)	0.602 (0.063)	0.574 (0.049)	0.575 (0.051)	0.548 (0.039)	0.552 (0.052)	0.554 (0.055)	0.630 (0.036)
0.3	lassojoint	0.989 (0.007)	0.989 (0.009)	0.988 (0.011)	0.944 (0.017)	0.944 (0.016)	0.943 (0.016)	0.942 (0.017)	0.942 (0.017)	0.941 (0.019)
0.5	naive	0.983 (0.005)	0.983 (0.005)	0.983 (0.005)	0.627 (0.007)	0.627 (0.007)	0.627 (0.007)	0.729 (0.006)	0.729 (0.006)	0.729 (0.006)
0.5	clust	0.622 (0.025)	0.637 (0.011)	0.555 (0.017)	0.579 (0.016)	0.576 (0.021)	0.531 (0.002)	0.555 (0.018)	0.570 (0.018)	0.653 (0.003)
0.5	strict-lassclust	0.620 (0.031)	0.624 (0.021)	0.561 (0.016)	0.559 (0.017)	0.557 (0.023)	0.534 (0.007)	0.643 (0.030)	0.583 (0.044)	0.577 (0.057)
0.5	non-strict-lassclust	0.625 (0.027)	0.642 (0.016)	0.550 (0.018)	0.598 (0.020)	0.579 (0.028)	0.527 (0.004)	0.535 (0.020)	0.562 (0.015)	0.655 (0.012)
0.5	lassojoint	0.993 (0.001)	0.993 (0.001)	0.993 (0.001)	0.943 (0.003)	0.943 (0.003)	0.943 (0.003)	0.943 (0.003)	0.943 (0.003)	0.943 (0.003)
0.8	naive	0.988 (0.000)	0.988 (0.000)	0.988 (0.000)	0.883 (0.005)	0.883 (0.005)	0.883 (0.005)	0.890 (0.005)	0.890 (0.005)	0.890 (0.005)
0.8	clust	0.580 (0.009)	0.575 (0.016)	0.602 (0.012)	0.564 (0.007)	0.553 (0.010)	0.477 (0.003)	0.519 (0.001)	0.525 (0.003)	0.642 (0.002)
0.8	strict-lassclust	0.586 (0.009)	0.596 (0.001)	0.605 (0.010)	0.573 (0.009)	0.568 (0.014)	0.479 (0.004)	0.441 (0.007)	0.338 (0.056)	0.645 (0.001)
0.8	non-strict-lassclust	0.581 (0.009)	0.579 (0.006)	0.603 (0.013)	0.567 (0.008)	0.547 (0.003)	0.478 (0.002)	0.547 (0.004)	0.548 (0.020)	0.643 (0.002)
0.8	lassojoint	0.974 (0.001)	0.982 (0.003)	0.974 (0.000)	0.574 (0.003)	0.618 (0.034)	0.576 (0.005)	0.692 (0.000)	0.715 (0.022)	0.693 (0.004)

Table 33. Summary Statistics for df = bank-marketing; SCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.890 (0.023)	0.890 (0.023)	0.890 (0.023)	0.111 (0.006)	0.111 (0.006)	0.111 (0.006)	0.200 (0.010)	0.200 (0.010)	0.200 (0.010)
0.3	clust	0.712 (0.052)	0.718 (0.049)	0.718 (0.049)	0.717 (0.013)	0.713 (0.047)	0.623 (0.043)	0.349 (0.034)	0.353 (0.035)	0.305 (0.021)
0.3	strict-lassclust	0.714 (0.049)	0.718 (0.050)	0.718 (0.050)	0.700 (0.106)	0.711 (0.062)	0.629 (0.028)	0.348 (0.037)	0.354 (0.031)	0.307 (0.019)
0.3	non-strict-lassclust	0.706 (0.056)	0.715 (0.053)	0.717 (0.050)	0.693 (0.117)	0.713 (0.065)	0.628 (0.034)	0.344 (0.045)	0.360 (0.065)	0.307 (0.020)
0.3	lassojoint	0.894 (0.032)	0.895 (0.032)	0.896 (0.023)	0.802 (0.239)	0.799 (0.241)	0.798 (0.243)	0.440 (0.109)	0.437 (0.110)	0.436 (0.110)
0.5	naive	0.910 (0.003)	0.910 (0.003)	0.910 (0.003)	0.407 (0.077)	0.407 (0.077)	0.407 (0.077)	0.265 (0.028)	0.265 (0.028)	0.265 (0.028)
0.5	clust	0.739 (0.038)	0.739 (0.038)	0.739 (0.038)	0.714 (0.007)	0.714 (0.007)	0.585 (0.104)	0.339 (0.011)	0.339 (0.011)	0.283 (0.026)
0.5	strict-lassclust	0.740 (0.037)	0.740 (0.038)	0.740 (0.038)	0.714 (0.007)	0.714 (0.007)	0.586 (0.099)	0.339 (0.011)	0.339 (0.011)	0.282 (0.026)
0.5	non-strict-lassclust	0.739 (0.038)	0.739 (0.038)	0.739 (0.038)	0.714 (0.007)	0.714 (0.007)	0.586 (0.099)	0.339 (0.011)	0.339 (0.011)	0.282 (0.026)
0.5	lassojoint	0.906 (0.006)	0.906 (0.006)	0.906 (0.006)	0.902 (0.003)	0.902 (0.004)	0.902 (0.004)	0.506 (0.029)	0.509 (0.027)	0.509 (0.028)
0.8	naive	0.920 (0.004)	0.920 (0.004)	0.920 (0.004)	0.909 (0.006)	0.909 (0.006)	0.909 (0.006)	0.547 (0.027)	0.547 (0.027)	0.547 (0.027)
0.8	clust	0.680 (0.022)	0.680 (0.022)	0.680 (0.022)	0.708 (0.006)	0.708 (0.006)	0.124 (0.033)	0.349 (0.016)	0.349 (0.016)	0.195 (0.010)
0.8	strict-lassclust	0.674 (0.031)	0.673 (0.030)	0.674 (0.033)	0.708 (0.006)	0.708 (0.006)	0.133 (0.039)	0.349 (0.016)	0.349 (0.016)	0.196 (0.011)
0.8	non-strict-lassclust	0.669 (0.022)	0.668 (0.020)	0.669 (0.023)	0.708 (0.006)	0.708 (0.006)	0.116 (0.021)	0.349 (0.016)	0.349 (0.016)	0.194 (0.007)
0.8	lassojoint	0.920 (0.005)	0.922 (0.005)	0.921 (0.004)	0.915 (0.004)	0.912 (0.006)	0.915 (0.004)	0.511 (0.031)	0.487 (0.055)	0.511 (0.032)

Table 34. Summary Statistics for df = banknote; SCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.997 (0.002)	0.997 (0.002)	0.997 (0.002)	0.462 (0.019)	0.462 (0.019)	0.462 (0.019)	0.624 (0.018)	0.624 (0.018)	0.624 (0.018)
0.3	clust	0.671 (0.057)	0.669 (0.054)	0.680 (0.040)	0.606 (0.077)	0.588 (0.067)	0.586 (0.059)	0.547 (0.077)	0.539 (0.073)	0.578 (0.033)
0.3	strict-lassclust	0.708 (0.110)	0.678 (0.073)	0.677 (0.038)	0.583 (0.061)	0.580 (0.051)	0.580 (0.050)	0.579 (0.127)	0.560 (0.085)	0.575 (0.030)
0.3	non-strict-lassclust	0.663 (0.053)	0.665 (0.050)	0.677 (0.038)	0.577 (0.074)	0.573 (0.060)	0.580 (0.050)	0.562 (0.056)	0.547 (0.057)	0.574 (0.030)
0.3	lassojoint	0.999 (0.001)	0.999 (0.001)	0.999 (0.001)	0.985 (0.007)	0.985 (0.007)	0.985 (0.007)	0.983 (0.008)	0.983 (0.008)	0.983 (0.008)
0.5	naive	0.997 (0.001)	0.997 (0.001)	0.997 (0.001)	0.723 (0.013)	0.723 (0.013)	0.723 (0.013)	0.757 (0.009)	0.757 (0.009)	0.757 (0.009)
0.5	clust	0.666 (0.020)	0.658 (0.029)	0.663 (0.028)	0.607 (0.029)	0.590 (0.031)	0.565 (0.023)	0.535 (0.015)	0.537 (0.027)	0.575 (0.018)
0.5	strict-lassclust	0.665 (0.022)	0.656 (0.029)	0.660 (0.028)	0.608 (0.031)	0.587 (0.031)	0.561 (0.021)	0.535 (0.018)	0.535 (0.027)	0.575 (0.018)
0.5	non-strict-lassclust	0.665 (0.022)	0.655 (0.029)	0.660 (0.028)	0.605 (0.030)	0.586 (0.030)	0.561 (0.021)	0.534 (0.018)	0.535 (0.027)	0.575 (0.018)
0.5	lassojoint	0.999 (0.000)	0.999 (0.000)	0.999 (0.000)	0.982 (0.006)	0.982 (0.006)	0.982 (0.006)	0.980 (0.007)	0.980 (0.007)	0.980 (0.007)
0.8	naive	0.996 (0.001)	0.996 (0.001)	0.996 (0.001)	0.949 (0.000)	0.949 (0.000)	0.949 (0.000)	0.945 (0.000)	0.945 (0.000)	0.945 (0.000)
0.8	clust	0.651 (0.027)	0.688 (0.017)	0.687 (0.018)	0.590 (0.032)	0.598 (0.022)	0.522 (0.024)	0.563 (0.019)	0.588 (0.010)	0.611 (0.005)
0.8	strict-lassclust	0.809 (0.005)	0.681 (0.015)	0.681 (0.014)	0.562 (0.031)	0.593 (0.020)	0.529 (0.015)	0.714 (0.014)	0.584 (0.008)	0.618 (0.001)
0.8	non-strict-lassclust	0.622 (0.028)	0.681 (0.015)	0.681 (0.014)	0.553 (0.033)	0.593 (0.020)	0.529 (0.015)	0.612 (0.004)	0.584 (0.008)	0.618 (0.001)
0.8	lassojoint	1.000 (0.000)	1.000 (0.000)	1.000 (0.000)	0.988 (0.000)	0.981 (0.002)	0.981 (0.002)	0.986 (0.001)	0.978 (0.002)	0.978 (0.002)

Table 35. Summary Statistics for df = breastc; SCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.984 (0.011)	0.984 (0.011)	0.984 (0.011)	0.348 (0.031)	0.348 (0.031)	0.348 (0.031)	0.514 (0.034)	0.514 (0.034)	0.514 (0.034)
0.3	clust	0.975 (0.022)	0.976 (0.022)	0.979 (0.020)	0.956 (0.012)	0.957 (0.012)	0.964 (0.013)	0.935 (0.018)	0.936 (0.019)	0.947 (0.020)
0.3	strict-lassclust	0.995 (0.003)	0.994 (0.003)	0.994 (0.003)	0.962 (0.012)	0.962 (0.012)	0.966 (0.016)	0.944 (0.018)	0.944 (0.018)	0.951 (0.023)
0.3	non-strict-lassclust	0.995 (0.003)	0.994 (0.003)	0.994 (0.003)	0.962 (0.011)	0.963 (0.011)	0.966 (0.016)	0.944 (0.018)	0.945 (0.017)	0.951 (0.023)
0.3	lassojoint	0.982 (0.016)	0.983 (0.015)	0.983 (0.016)	0.879 (0.176)	0.868 (0.191)	0.868 (0.191)	0.875 (0.125)	0.867 (0.134)	0.867 (0.134)
0.5	naive	0.993 (0.003)	0.993 (0.003)	0.993 (0.003)	0.706 (0.032)	0.706 (0.032)	0.706 (0.032)	0.691 (0.024)	0.691 (0.024)	0.691 (0.024)
0.5	clust	0.991 (0.008)	0.991 (0.007)	0.988 (0.009)	0.966 (0.007)	0.970 (0.006)	0.963 (0.006)	0.949 (0.012)	0.955 (0.010)	0.946 (0.008)
0.5	strict-lassclust	0.997 (0.001)	0.996 (0.001)	0.995 (0.001)	0.978 (0.004)	0.978 (0.006)	0.965 (0.006)	0.968 (0.007)	0.967 (0.008)	0.949 (0.009)
0.5	non-strict-lassclust	0.997 (0.001)	0.996 (0.001)	0.995 (0.001)	0.978 (0.004)	0.978 (0.006)	0.965 (0.006)	0.968 (0.007)	0.967 (0.008)	0.949 (0.009)
0.5	lassojoint	0.996 (0.001)	0.996 (0.001)	0.996 (0.001)	0.958 (0.010)	0.958 (0.010)	0.958 (0.010)	0.940 (0.014)	0.940 (0.014)	0.940 (0.014)
0.8	naive	0.992 (0.003)	0.992 (0.003)	0.992 (0.003)	0.941 (0.015)	0.941 (0.015)	0.941 (0.015)	0.921 (0.016)	0.921 (0.016)	0.921 (0.016)
0.8	clust	0.962 (0.015)	0.961 (0.021)	0.962 (0.016)	0.957 (0.007)	0.959 (0.009)	0.940 (0.038)	0.936 (0.012)	0.941 (0.018)	0.919 (0.047)
0.8	strict-lassclust	0.992 (0.002)	0.991 (0.003)	0.988 (0.004)	0.963 (0.006)	0.967 (0.005)	0.867 (0.083)	0.945 (0.009)	0.953 (0.008)	0.841 (0.078)
0.8	non-strict-lassclust	0.992 (0.002)	0.991 (0.003)	0.988 (0.004)	0.963 (0.006)	0.967 (0.005)	0.867 (0.083)	0.945 (0.009)	0.953 (0.008)	0.841 (0.078)
0.8	lassojoint	0.993 (0.002)	0.993 (0.002)	0.993 (0.002)	0.964 (0.005)	0.934 (0.134)	0.934 (0.134)	0.950 (0.005)	0.929 (0.094)	0.929 (0.094)

Table 36. Summary Statistics for df = credit-a; SCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.813 (0.030)	0.813 (0.030)	0.813 (0.030)	0.457 (0.029)	0.457 (0.029)	0.457 (0.029)	0.626 (0.027)	0.626 (0.027)	0.626 (0.027)
0.3	clust	0.741 (0.055)	0.742 (0.055)	0.752 (0.048)	0.637 (0.033)	0.638 (0.032)	0.659 (0.029)	0.400 (0.062)	0.403 (0.061)	0.477 (0.064)
0.3	strict-lassclust	0.756 (0.033)	0.758 (0.034)	0.762 (0.033)	0.636 (0.034)	0.637 (0.033)	0.652 (0.031)	0.393 (0.066)	0.396 (0.064)	0.451 (0.065)
0.3	non-strict-lassclust	0.756 (0.033)	0.758 (0.034)	0.762 (0.033)	0.636 (0.034)	0.637 (0.033)	0.652 (0.031)	0.392 (0.064)	0.395 (0.063)	0.451 (0.065)
0.3	lassojoint	0.808 (0.033)	0.807 (0.033)	0.808 (0.032)	0.570 (0.137)	0.573 (0.138)	0.575 (0.139)	0.642 (0.046)	0.642 (0.046)	0.643 (0.048)
0.5	naive	0.800 (0.017)	0.800 (0.017)	0.800 (0.017)	0.450 (0.007)	0.450 (0.007)	0.450 (0.007)	0.618 (0.007)	0.618 (0.007)	0.618 (0.007)
0.5	clust	0.745 (0.005)	0.747 (0.006)	0.750 (0.006)	0.621 (0.004)	0.621 (0.002)	0.648 (0.012)	0.312 (0.013)	0.313 (0.008)	0.425 (0.057)
0.5	strict-lassclust	0.751 (0.006)	0.753 (0.007)	0.756 (0.007)	0.612 (0.006)	0.613 (0.007)	0.634 (0.009)	0.276 (0.015)	0.276 (0.017)	0.356 (0.047)
0.5	non-strict-lassclust	0.751 (0.006)	0.753 (0.007)	0.756 (0.007)	0.612 (0.006)	0.613 (0.007)	0.634 (0.009)	0.276 (0.015)	0.276 (0.017)	0.356 (0.047)
0.5	lassojoint	0.800 (0.018)	0.800 (0.017)	0.804 (0.015)	0.481 (0.093)	0.480 (0.092)	0.480 (0.092)	0.620 (0.008)	0.620 (0.008)	0.620 (0.008)
0.8	naive	0.837 (0.000)	0.837 (0.000)	0.837 (0.000)	0.764 (0.000)	0.764 (0.000)	0.764 (0.000)	0.744 (0.000)	0.744 (0.000)	0.744 (0.000)
0.8	clust	0.794 (0.000)	0.797 (0.000)	0.804 (0.000)	0.636 (0.000)	0.636 (0.000)	0.744 (0.000)	0.403 (0.000)	0.403 (0.000)	0.719 (0.000)
0.8	strict-lassclust	0.794 (0.000)	0.798 (0.000)	0.807 (0.000)	0.631 (0.000)	0.631 (0.000)	0.733 (0.000)	0.379 (0.000)	0.379 (0.000)	0.683 (0.000)
0.8	non-strict-lassclust	0.794 (0.000)	0.798 (0.000)	0.807 (0.000)	0.631 (0.000)	0.631 (0.000)	0.733 (0.000)	0.379 (0.000)	0.379 (0.000)	0.683 (0.000)
0.8	lassojoint	0.835 (0.000)	0.835 (0.000)	0.835 (0.000)	0.713 (0.000)	0.713 (0.000)	0.713 (0.000)	0.743 (0.000)	0.743 (0.000)	0.743 (0.000)

Table 37. Summary Statistics for df = credit-g; SCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.679 (0.042)	0.679 (0.042)	0.679 (0.042)	0.299 (0.023)	0.299 (0.023)	0.299 (0.023)	0.460 (0.027)	0.460 (0.027)	0.460 (0.027)
0.3	clust	0.617 (0.045)	0.618 (0.041)	0.619 (0.034)	0.648 (0.107)	0.667 (0.084)	0.631 (0.096)	0.370 (0.049)	0.371 (0.044)	0.414 (0.040)
0.3	strict-lassclust	0.613 (0.042)	0.619 (0.035)	0.617 (0.034)	0.646 (0.111)	0.659 (0.099)	0.637 (0.100)	0.381 (0.094)	0.367 (0.050)	0.409 (0.056)
0.3	non-strict-lassclust	0.617 (0.049)	0.617 (0.043)	0.619 (0.034)	0.655 (0.103)	0.670 (0.081)	0.636 (0.103)	0.353 (0.050)	0.360 (0.041)	0.403 (0.053)
0.3	lassojoint	0.662 (0.063)	0.661 (0.066)	0.652 (0.074)	0.596 (0.178)	0.596 (0.179)	0.596 (0.179)	0.420 (0.133)	0.417 (0.138)	0.436 (0.147)
0.5	naive	0.748 (0.011)	0.748 (0.011)	0.748 (0.011)	0.387 (0.029)	0.387 (0.029)	0.387 (0.029)	0.476 (0.017)	0.476 (0.017)	0.476 (0.017)
0.5	clust	0.621 (0.023)	0.666 (0.038)	0.639 (0.012)	0.729 (0.020)	0.733 (0.018)	0.636 (0.021)	0.448 (0.060)	0.430 (0.033)	0.457 (0.020)
0.5	strict-lassclust	0.630 (0.021)	0.676 (0.028)	0.643 (0.013)	0.739 (0.022)	0.739 (0.016)	0.653 (0.026)	0.444 (0.070)	0.424 (0.040)	0.451 (0.021)
0.5	non-strict-lassclust	0.624 (0.020)	0.666 (0.037)	0.644 (0.010)	0.734 (0.022)	0.735 (0.017)	0.656 (0.021)	0.449 (0.063)	0.432 (0.042)	0.452 (0.023)
0.5	lassojoint	0.741 (0.018)	0.741 (0.018)	0.741 (0.018)	0.349 (0.148)	0.349 (0.148)	0.349 (0.148)	0.424 (0.068)	0.424 (0.068)	0.424 (0.068)
0.8	naive	0.751 (0.022)	0.751 (0.022)	0.751 (0.022)	0.754 (0.017)	0.754 (0.017)	0.754 (0.017)	0.584 (0.002)	0.584 (0.002)	0.584 (0.002)
0.8	clust	0.579 (0.023)	0.586 (0.023)	0.592 (0.022)	0.684 (0.014)	0.687 (0.007)	0.366 (0.028)	0.350 (0.023)	0.343 (0.019)	0.454 (0.011)
0.8	strict-lassclust	0.602 (0.033)	0.609 (0.033)	0.623 (0.037)	0.697 (0.017)	0.698 (0.015)	0.454 (0.075)	0.341 (0.017)	0.342 (0.016)	0.469 (0.021)
0.8	non-strict-lassclust	0.588 (0.029)	0.596 (0.021)	0.611 (0.026)	0.682 (0.015)	0.690 (0.007)	0.418 (0.040)	0.344 (0.020)	0.345 (0.019)	0.464 (0.016)
0.8	lassojoint	0.731 (0.006)	0.731 (0.006)	0.735 (0.010)	0.725 (0.002)	0.725 (0.002)	0.727 (0.004)	0.226 (0.016)	0.226 (0.016)	0.220 (0.010)

Table 38. Summary Statistics for df = dhfr; SCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.592 (0.074)	0.592 (0.074)	0.592 (0.074)	0.534 (0.083)	0.534 (0.083)	0.534 (0.083)	0.559 (0.068)	0.559 (0.068)	0.559 (0.068)
0.3	clust	0.639 (0.081)	0.603 (0.075)	0.521 (0.117)	0.505 (0.177)	0.505 (0.149)	0.504 (0.164)	0.438 (0.090)	0.436 (0.079)	0.475 (0.036)
0.3	strict-lassclust	0.617 (0.058)	0.628 (0.062)	0.622 (0.072)	0.547 (0.155)	0.571 (0.105)	0.566 (0.092)	0.645 (0.125)	0.635 (0.139)	0.525 (0.122)
0.3	non-strict-lassclust	0.688 (0.066)	0.703 (0.081)	0.614 (0.077)	0.637 (0.100)	0.631 (0.113)	0.534 (0.094)	0.509 (0.178)	0.492 (0.200)	0.486 (0.081)
0.3	lassojoint	0.866 (0.047)	0.831 (0.121)	0.837 (0.095)	0.659 (0.174)	0.585 (0.200)	0.643 (0.175)	0.675 (0.095)	0.640 (0.097)	0.646 (0.122)
0.5	naive	0.647 (0.012)	0.647 (0.012)	0.647 (0.012)	0.596 (0.014)	0.596 (0.014)	0.596 (0.014)	0.631 (0.003)	0.631 (0.003)	0.631 (0.003)
0.5	clust	0.531 (0.023)	0.525 (0.010)	0.476 (0.009)	0.497 (0.047)	0.481 (0.013)	0.443 (0.023)	0.430 (0.023)	0.423 (0.039)	0.465 (0.038)
0.5	strict-lassclust	0.603 (0.045)	0.545 (0.073)	0.685 (0.022)	0.640 (0.021)	0.591 (0.028)	0.460 (0.013)	0.265 (0.271)	0.680 (0.123)	0.546 (0.028)
0.5	non-strict-lassclust	0.795 (0.081)	0.808 (0.106)	0.685 (0.032)	0.725 (0.179)	0.681 (0.098)	0.527 (0.056)	0.712 (0.117)	0.654 (0.115)	0.516 (0.004)
0.5	lassojoint	0.951 (0.015)	0.954 (0.022)	0.954 (0.027)	0.710 (0.130)	0.748 (0.077)	0.783 (0.048)	0.737 (0.074)	0.755 (0.053)	0.776 (0.027)
0.8	naive	0.652 (0.000)	0.652 (0.000)	0.652 (0.000)	0.619 (0.000)	0.619 (0.000)	0.619 (0.000)	0.584 (0.000)	0.584 (0.000)	0.584 (0.000)
0.8	clust	0.556 (0.000)	0.522 (0.000)	0.481 (0.000)	0.474 (0.000)	0.526 (0.000)	0.443 (0.000)	0.354 (0.000)	0.410 (0.000)	0.426 (0.000)
0.8	strict-lassclust	0.500 (0.000)	0.658 (0.000)	0.658 (0.000)	0.649 (0.000)	0.567 (0.000)	0.412 (0.000)	0.788 (0.000)	0.512 (0.000)	0.513 (0.000)
0.8	non-strict-lassclust	0.710 (0.000)	0.661 (0.000)	0.658 (0.000)	0.722 (0.000)	0.567 (0.000)	0.423 (0.000)	0.597 (0.000)	0.488 (0.000)	0.517 (0.000)
0.8	lassojoint	0.929 (0.000)	0.943 (0.000)	0.929 (0.000)	0.660 (0.000)	0.887 (0.000)	0.567 (0.000)	0.667 (0.000)	0.836 (0.000)	0.611 (0.000)

Table 39. Summary Statistics for df = diabetes; SCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.794 (0.032)	0.794 (0.032)	0.794 (0.032)	0.352 (0.030)	0.352 (0.030)	0.352 (0.030)	0.516 (0.032)	0.516 (0.032)	0.516 (0.032)
0.3	clust	0.729 (0.046)	0.731 (0.047)	0.727 (0.049)	0.684 (0.033)	0.682 (0.038)	0.626 (0.055)	0.570 (0.055)	0.572 (0.057)	0.595 (0.049)
0.3	strict-lassclust	0.725 (0.046)	0.727 (0.045)	0.725 (0.050)	0.686 (0.031)	0.682 (0.030)	0.634 (0.058)	0.541 (0.132)	0.526 (0.139)	0.588 (0.055)
0.3	non-strict-lassclust	0.732 (0.045)	0.733 (0.047)	0.731 (0.049)	0.674 (0.048)	0.671 (0.048)	0.610 (0.070)	0.578 (0.054)	0.581 (0.053)	0.591 (0.047)
0.3	lassojoint	0.796 (0.043)	0.800 (0.032)	0.800 (0.033)	0.682 (0.128)	0.685 (0.124)	0.684 (0.129)	0.593 (0.063)	0.597 (0.061)	0.595 (0.062)
0.5	naive	0.791 (0.014)	0.791 (0.014)	0.791 (0.014)	0.533 (0.037)	0.533 (0.037)	0.533 (0.037)	0.588 (0.025)	0.588 (0.025)	0.588 (0.025)
0.5	clust	0.733 (0.046)	0.730 (0.027)	0.723 (0.022)	0.699 (0.036)	0.684 (0.023)	0.551 (0.067)	0.584 (0.066)	0.563 (0.048)	0.594 (0.033)
0.5	strict-lassclust	0.682 (0.041)	0.720 (0.026)	0.723 (0.021)	0.653 (0.027)	0.678 (0.023)	0.539 (0.066)	0.653 (0.137)	0.543 (0.083)	0.590 (0.033)
0.5	non-strict-lassclust	0.738 (0.049)	0.734 (0.024)	0.725 (0.020)	0.669 (0.024)	0.687 (0.019)	0.535 (0.068)	0.599 (0.074)	0.610 (0.039)	0.588 (0.033)
0.5	lassojoint	0.791 (0.015)	0.791 (0.015)	0.791 (0.015)	0.707 (0.041)	0.705 (0.039)	0.707 (0.041)	0.639 (0.012)	0.639 (0.013)	0.639 (0.012)
0.8	naive	0.797 (0.019)	0.797 (0.019)	0.797 (0.019)	0.724 (0.016)	0.724 (0.016)	0.724 (0.016)	0.653 (0.016)	0.653 (0.016)	0.653 (0.016)
0.8	clust	0.732 (0.013)	0.716 (0.037)	0.712 (0.037)	0.678 (0.015)	0.666 (0.032)	0.346 (0.008)	0.532 (0.009)	0.512 (0.042)	0.507 (0.014)
0.8	strict-lassclust	0.725 (0.010)	0.709 (0.042)	0.703 (0.042)	0.675 (0.033)	0.348 (0.005)	0.193 (0.202)	0.499 (0.023)	0.510 (0.015)	0.510 (0.015)
0.8	non-strict-lassclust	0.731 (0.012)	0.712 (0.045)	0.707 (0.047)	0.675 (0.018)	0.672 (0.030)	0.344 (0.011)	0.592 (0.048)	0.534 (0.061)	0.509 (0.016)
0.8	lassojoint	0.798 (0.015)	0.798 (0.015)	0.798 (0.015)	0.353 (0.016)	0.353 (0.016)	0.352 (0.016)	0.512 (0.016)	0.512 (0.016)	0.512 (0.016)

Table 40. Summary Statistics for df = heart-c; SCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.708 (0.065)	0.708 (0.065)	0.708 (0.065)	0.466 (0.040)	0.466 (0.040)	0.466 (0.040)	0.633 (0.038)	0.633 (0.038)	0.633 (0.038)
0.3	clust	0.696 (0.064)	0.706 (0.058)	0.710 (0.055)	0.630 (0.096)	0.656 (0.064)	0.603 (0.062)	0.621 (0.127)	0.664 (0.062)	0.674 (0.058)
0.3	strict-lassclust	0.696 (0.062)	0.705 (0.058)	0.708 (0.055)	0.627 (0.099)	0.653 (0.067)	0.600 (0.060)	0.611 (0.139)	0.661 (0.064)	0.673 (0.058)
0.3	non-strict-lassclust	0.696 (0.061)	0.704 (0.058)	0.708 (0.055)	0.626 (0.100)	0.654 (0.065)	0.600 (0.060)	0.606 (0.150)	0.662 (0.064)	0.673 (0.058)
0.3	lassojoint	0.671 (0.101)	0.678 (0.103)	0.677 (0.102)	0.554 (0.091)	0.554 (0.096)	0.563 (0.096)	0.616 (0.098)	0.617 (0.099)	0.608 (0.112)
0.5	naive	0.739 (0.003)	0.739 (0.003)	0.739 (0.003)	0.433 (0.003)	0.433 (0.003)	0.433 (0.003)	0.605 (0.003)	0.605 (0.003)	0.605 (0.003)
0.5	clust	0.715 (0.003)	0.725 (0.002)	0.731 (0.005)	0.600 (0.001)	0.600 (0.001)	0.456 (0.000)	0.640 (0.002)	0.647 (0.001)	0.608 (0.002)
0.5	strict-lassclust	0.714 (0.003)	0.723 (0.002)	0.731 (0.005)	0.600 (0.001)	0.589 (0.001)	0.444 (0.000)	0.640 (0.001)	0.641 (0.002)	0.603 (0.002)
0.5	non-strict-lassclust	0.714 (0.003)	0.723 (0.002)	0.731 (0.005)	0.600 (0.001)	0.589 (0.001)	0.444 (0.000)	0.640 (0.001)	0.641 (0.002)	0.603 (0.002)
0.5	lassojoint	0.727 (0.006)	0.735 (0.005)	0.727 (0.006)	0.686 (0.029)	0.598 (0.020)	0.686 (0.029)	0.631 (0.006)	0.617 (0.005)	0.631 (0.006)
0.8	naive	0.746 (0.000)	0.746 (0.000)	0.746 (0.000)	0.611 (0.002)	0.611 (0.002)	0.611 (0.002)	0.599 (0.007)	0.599 (0.007)	0.599 (0.007)
0.8	clust	0.773 (0.003)	0.784 (0.004)	0.792 (0.005)	0.711 (0.001)	0.744 (0.002)	0.400 (0.003)	0.650 (0.005)	0.701 (0.002)	0.571 (0.003)
0.8	strict-lassclust	0.773 (0.003)	0.782 (0.004)	0.793 (0.005)	0.733 (0.003)	0.733 (0.001)	0.400 (0.003)	0.684 (0.002)	0.692 (0.003)	0.571 (0.003)
0.8	non-strict-lassclust	0.773 (0.003)	0.782 (0.004)	0.793 (0.005)	0.733 (0.003)	0.733 (0.001)	0.400 (0.003)	0.684 (0.002)	0.692 (0.003)	0.571 (0.003)
0.8	lassojoint	0.777 (0.003)	0.777 (0.003)	0.777 (0.003)	0.777 (0.011)	0.777 (0.011)	0.777 (0.011)	0.696 (0.014)	0.696 (0.014)	0.696 (0.014)

Table 41. Summary Statistics for df = qsar-biodeg; SCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.815 (0.031)	0.815 (0.031)	0.815 (0.031)	0.372 (0.025)	0.372 (0.025)	0.372 (0.025)	0.516 (0.024)	0.516 (0.024)	0.516 (0.024)
0.3	clust	0.785 (0.026)	0.784 (0.026)	0.778 (0.028)	0.668 (0.032)	0.661 (0.033)	0.546 (0.044)	0.632 (0.031)	0.629 (0.030)	0.583 (0.030)
0.3	strict-lassclust	0.783 (0.030)	0.783 (0.029)	0.781 (0.030)	0.643 (0.076)	0.642 (0.069)	0.543 (0.090)	0.583 (0.133)	0.584 (0.120)	0.551 (0.109)
0.3	non-strict-lassclust	0.777 (0.052)	0.777 (0.045)	0.773 (0.039)	0.618 (0.114)	0.614 (0.111)	0.506 (0.076)	0.604 (0.055)	0.606 (0.052)	0.568 (0.038)
0.3	lassojoint	0.837 (0.063)	0.845 (0.039)	0.841 (0.052)	0.754 (0.094)	0.759 (0.089)	0.746 (0.104)	0.679 (0.059)	0.678 (0.053)	0.674 (0.058)
0.5	naive	0.866 (0.008)	0.866 (0.008)	0.866 (0.008)	0.512 (0.026)	0.512 (0.026)	0.512 (0.026)	0.583 (0.015)	0.583 (0.015)	0.583 (0.015)
0.5	clust	0.807 (0.015)	0.802 (0.018)	0.798 (0.020)	0.722 (0.017)	0.709 (0.020)	0.582 (0.021)	0.678 (0.014)	0.673 (0.014)	0.617 (0.016)
0.5	strict-lassclust	0.814 (0.007)	0.809 (0.024)	0.818 (0.014)	0.699 (0.032)	0.677 (0.021)	0.626 (0.058)	0.658 (0.016)	0.654 (0.008)	0.633 (0.033)
0.5	non-strict-lassclust	0.801 (0.042)	0.817 (0.017)	0.772 (0.046)	0.691 (0.056)	0.715 (0.066)	0.466 (0.088)	0.663 (0.031)	0.661 (0.037)	0.566 (0.034)
0.5	lassojoint	0.874 (0.014)	0.878 (0.018)	0.877 (0.013)	0.723 (0.181)	0.832 (0.020)	0.832 (0.013)	0.702 (0.097)	0.751 (0.031)	0.753 (0.025)
0.8	naive	0.904 (0.000)	0.904 (0.000)	0.904 (0.000)	0.807 (0.000)	0.807 (0.000)	0.807 (0.000)	0.775 (0.000)	0.775 (0.000)	0.775 (0.000)
0.8	clust	0.827 (0.000)	0.811 (0.000)	0.807 (0.000)	0.728 (0.000)	0.684 (0.000)	0.522 (0.000)	0.695 (0.000)	0.677 (0.000)	0.612 (0.000)
0.8	strict-lassclust	0.802 (0.000)	0.813 (0.000)	0.806 (0.000)	0.633 (0.000)	0.649 (0.000)	0.443 (0.000)	0.659 (0.000)	0.661 (0.000)	0.577 (0.000)
0.8	non-strict-lassclust	0.835 (0.000)	0.813 (0.000)	0.806 (0.000)	0.769 (0.000)	0.658 (0.000)	0.440 (0.000)	0.700 (0.000)	0.667 (0.000)	0.576 (0.000)
0.8	lassojoint	0.917 (0.000)	0.917 (0.000)	0.916 (0.000)	0.851 (0.000)	0.851 (0.000)	0.851 (0.000)	0.798 (0.000)	0.798 (0.000)	0.795 (0.000)

Table 42. Summary Statistics for df = spambase; SCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.936 (0.010)	0.936 (0.010)	0.936 (0.010)	0.397 (0.011)	0.397 (0.011)	0.397 (0.011)	0.565 (0.011)	0.565 (0.011)	0.565 (0.011)
0.3	clust	0.888 (0.053)	0.891 (0.041)	0.896 (0.068)	0.789 (0.136)	0.814 (0.110)	0.803 (0.148)	0.757 (0.078)	0.776 (0.059)	0.787 (0.085)
0.3	strict-lassclust	0.889 (0.043)	0.898 (0.032)	0.906 (0.028)	0.778 (0.136)	0.803 (0.106)	0.792 (0.158)	0.713 (0.162)	0.724 (0.158)	0.783 (0.089)
0.3	non-strict-lassclust	0.911 (0.012)	0.913 (0.010)	0.912 (0.020)	0.790 (0.133)	0.816 (0.103)	0.794 (0.156)	0.762 (0.071)	0.776 (0.056)	0.783 (0.088)
0.3	lassojoint	0.946 (0.011)	0.946 (0.011)	0.946 (0.012)	0.872 (0.100)	0.864 (0.119)	0.863 (0.119)	0.852 (0.063)	0.848 (0.073)	0.846 (0.074)
0.5	naive	0.956 (0.004)	0.956 (0.004)	0.956 (0.004)	0.530 (0.025)	0.530 (0.025)	0.530 (0.025)	0.631 (0.012)	0.631 (0.012)	0.631 (0.012)
0.5	clust	0.894 (0.062)	0.903 (0.046)	0.918 (0.036)	0.831 (0.022)	0.840 (0.012)	0.864 (0.021)	0.757 (0.046)	0.778 (0.024)	0.828 (0.036)
0.5	strict-lassclust	0.907 (0.006)	0.903 (0.018)	0.916 (0.005)	0.834 (0.012)	0.800 (0.085)	0.855 (0.027)	0.766 (0.021)	0.658 (0.260)	0.827 (0.022)
0.5	non-strict-lassclust	0.908 (0.006)	0.911 (0.005)	0.916 (0.006)	0.833 (0.012)	0.835 (0.011)	0.856 (0.028)	0.764 (0.022)	0.773 (0.022)	0.827 (0.022)
0.5	lassojoint	0.967 (0.004)	0.967 (0.004)	0.968 (0.004)	0.919 (0.006)	0.917 (0.006)	0.918 (0.006)	0.898 (0.008)	0.897 (0.008)	0.897 (0.008)
0.8	naive	0.963 (0.001)	0.963 (0.001)	0.963 (0.001)	0.913 (0.003)	0.913 (0.003)	0.913 (0.003)	0.894 (0.003)	0.894 (0.003)	0.894 (0.003)
0.8	clust	0.886 (0.045)	0.854 (0.055)	0.851 (0.053)	0.703 (0.062)	0.834 (0.063)	0.724 (0.088)	0.713 (0.027)	0.790 (0.032)	0.737 (0.049)
0.8	strict-lassclust	0.817 (0.038)	0.877 (0.045)	0.907 (0.018)	0.751 (0.055)	0.801 (0.080)	0.619 (0.047)	0.689 (0.027)	0.709 (0.129)	0.671 (0.021)
0.8	non-strict-lassclust	0.894 (0.040)	0.897 (0.042)	0.910 (0.004)	0.703 (0.061)	0.834 (0.063)	0.620 (0.048)	0.715 (0.026)	0.790 (0.033)	0.672 (0.022)
0.8	lassojoint	0.942 (0.004)	0.942 (0.004)	0.942 (0.004)	0.402 (0.004)	0.402 (0.003)	0.402 (0.003)	0.572 (0.003)	0.572 (0.003)	0.572 (0.003)

Table 43. Summary Statistics for df = wdbc; SCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.959 (0.022)	0.959 (0.022)	0.959 (0.022)	0.435 (0.037)	0.435 (0.037)	0.435 (0.037)	0.566 (0.029)	0.566 (0.029)	0.566 (0.029)
0.3	clust	0.900 (0.024)	0.901 (0.024)	0.899 (0.028)	0.891 (0.024)	0.892 (0.024)	0.886 (0.028)	0.842 (0.036)	0.844 (0.036)	0.840 (0.039)
0.3	strict-lassclust	0.960 (0.017)	0.959 (0.014)	0.964 (0.011)	0.865 (0.055)	0.860 (0.054)	0.903 (0.026)	0.775 (0.141)	0.767 (0.131)	0.861 (0.046)
0.3	non-strict-lassclust	0.963 (0.013)	0.964 (0.013)	0.965 (0.012)	0.904 (0.021)	0.907 (0.020)	0.908 (0.018)	0.867 (0.030)	0.871 (0.028)	0.872 (0.025)
0.3	lassojoint	0.973 (0.020)	0.973 (0.020)	0.974 (0.021)	0.883 (0.111)	0.880 (0.121)	0.879 (0.127)	0.867 (0.082)	0.866 (0.088)	0.867 (0.092)
0.5	naive	0.970 (0.006)	0.970 (0.006)	0.970 (0.006)	0.684 (0.049)	0.684 (0.049)	0.684 (0.049)	0.731 (0.044)	0.731 (0.044)	0.731 (0.044)
0.5	clust	0.899 (0.008)	0.901 (0.008)	0.911 (0.004)	0.890 (0.016)	0.890 (0.009)	0.889 (0.007)	0.864 (0.025)	0.866 (0.014)	0.874 (0.016)
0.5	strict-lassclust	0.950 (0.009)	0.950 (0.004)	0.957 (0.004)	0.856 (0.042)	0.860 (0.027)	0.900 (0.020)	0.806 (0.080)	0.817 (0.045)	0.885 (0.028)
0.5	non-strict-lassclust	0.952 (0.003)	0.947 (0.006)	0.957 (0.004)	0.875 (0.006)	0.885 (0.005)	0.902 (0.020)	0.845 (0.008)	0.861 (0.007)	0.890 (0.022)
0.5	lassojoint	0.981 (0.006)	0.980 (0.004)	0.977 (0.003)	0.846 (0.201)	0.927 (0.015)	0.926 (0.015)	0.865 (0.137)	0.917 (0.021)	0.915 (0.021)
0.8	naive	0.983 (0.000)	0.983 (0.000)	0.983 (0.000)	0.912 (0.000)	0.912 (0.000)	0.912 (0.000)	0.887 (0.000)	0.887 (0.000)	0.887 (0.000)
0.8	clust	0.911 (0.000)	0.927 (0.000)	0.949 (0.000)	0.906 (0.000)	0.906 (0.000)	0.912 (0.000)	0.857 (0.000)	0.862 (0.000)	0.885 (0.000)
0.8	strict-lassclust	0.974 (0.000)	0.948 (0.000)	0.980 (0.000)	0.888 (0.000)	0.694 (0.000)	0.912 (0.000)	0.826 (0.000)	0.297 (0.000)	0.891 (0.000)
0.8	non-strict-lassclust	0.980 (0.000)	0.985 (0.000)	0.980 (0.000)	0.929 (0.000)	0.718 (0.000)	0.906 (0.000)	0.898 (0.000)	0.721 (0.000)	0.884 (0.000)
0.8	lassojoint	0.987 (0.000)	0.981 (0.000)	0.981 (0.000)	0.524 (0.000)	0.959 (0.000)	0.959 (0.000)	0.609 (0.000)	0.945 (0.000)	0.945 (0.000)

Table 44. Summary Statistics for df = wine-quality; SCAR scheme

calc-c	method	AUC			ACC			F1		
		q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1	q=0.25	q=0.5	q=1
0.3	naive	0.814 (0.046)	0.814 (0.046)	0.814 (0.046)	0.137 (0.013)	0.137 (0.013)	0.137 (0.013)	0.238 (0.019)	0.238 (0.019)	0.238 (0.019)
0.3	clust	0.689 (0.043)	0.693 (0.037)	0.694 (0.034)	0.620 (0.080)	0.621 (0.088)	0.445 (0.084)	0.342 (0.062)	0.343 (0.071)	0.280 (0.050)
0.3	strict-lassclust	0.703 (0.044)	0.702 (0.038)	0.695 (0.035)	0.604 (0.162)	0.606 (0.129)	0.460 (0.086)	0.342 (0.093)	0.339 (0.075)	0.283 (0.054)
0.3	non-strict-lassclust	0.686 (0.046)	0.690 (0.042)	0.696 (0.033)	0.627 (0.112)	0.636 (0.092)	0.458 (0.082)	0.334 (0.073)	0.337 (0.079)	0.283 (0.052)
0.3	lassojoint	0.804 (0.090)	0.805 (0.086)	0.807 (0.081)	0.685 (0.293)	0.724 (0.266)	0.710 (0.275)	0.381 (0.155)	0.385 (0.155)	0.376 (0.147)
0.5	naive	0.829 (0.013)	0.829 (0.013)	0.829 (0.013)	0.448 (0.034)	0.448 (0.034)	0.448 (0.034)	0.320 (0.019)	0.320 (0.019)	0.320 (0.019)
0.5	clust	0.705 (0.027)	0.703 (0.023)	0.701 (0.024)	0.622 (0.044)	0.620 (0.045)	0.279 (0.073)	0.346 (0.024)	0.346 (0.027)	0.257 (0.017)
0.5	strict-lassclust	0.711 (0.021)	0.706 (0.018)	0.707 (0.017)	0.635 (0.042)	0.641 (0.027)	0.290 (0.070)	0.358 (0.024)	0.352 (0.022)	0.264 (0.013)
0.5	non-strict-lassclust	0.704 (0.026)	0.704 (0.021)	0.701 (0.023)	0.636 (0.039)	0.627 (0.048)	0.295 (0.067)	0.348 (0.026)	0.344 (0.027)	0.260 (0.016)
0.5	lassojoint	0.827 (0.013)	0.821 (0.019)	0.824 (0.015)	0.839 (0.028)	0.836 (0.041)	0.818 (0.080)	0.404 (0.062)	0.393 (0.050)	0.398 (0.057)
0.8	naive	0.825 (0.003)	0.825 (0.003)	0.825 (0.003)	0.833 (0.000)	0.833 (0.000)	0.833 (0.000)	0.414 (0.007)	0.414 (0.007)	0.414 (0.007)
0.8	clust	0.704 (0.012)	0.689 (0.010)	0.674 (0.007)	0.589 (0.013)	0.608 (0.020)	0.161 (0.001)	0.333 (0.002)	0.344 (0.006)	0.254 (0.004)
0.8	strict-lassclust	0.729 (0.019)	0.686 (0.005)	0.674 (0.004)	0.697 (0.028)	0.658 (0.012)	0.149 (0.003)	0.387 (0.008)	0.354 (0.005)	0.251 (0.004)
0.8	non-strict-lassclust	0.696 (0.010)	0.679 (0.006)	0.673 (0.006)	0.533 (0.019)	0.515 (0.008)	0.148 (0.001)	0.323 (0.003)	0.306 (0.000)	0.251 (0.004)
0.8	lassojoint	0.819 (0.005)	0.819 (0.005)	0.830 (0.002)	0.847 (0.003)	0.847 (0.003)	0.852 (0.005)	0.362 (0.011)	0.362 (0.011)	0.381 (0.017)

2.2 SCAR scheme - with SMOTE algorithm; pecking_part = 1

Table 45. Summary Statistics for df = adult; SCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.797 (0.003)	0.797 (0.003)	0.241 (0.005)	0.527 (0.014)	0.388 (0.006)	0.490 (0.010)
0.3	clust	0.763 (0.005)	0.770 (0.004)	0.406 (0.011)	0.746 (0.003)	0.442 (0.008)	0.507 (0.010)
0.3	strict-lassclust	0.763 (0.005)	0.770 (0.004)	0.405 (0.011)	0.746 (0.003)	0.442 (0.008)	0.507 (0.010)
0.3	non-strict-lassclust	0.763 (0.005)	0.770 (0.004)	0.405 (0.011)	0.746 (0.003)	0.442 (0.008)	0.507 (0.010)
0.3	lassojoint	0.795 (0.003)	0.794 (0.003)	0.792 (0.003)	0.785 (0.004)	0.480 (0.033)	0.306 (0.031)
0.3	spy	0.744 (0.006)		0.759 (0.005)		0.491 (0.011)	
0.5	naive	0.804 (0.001)	0.804 (0.001)	0.335 (0.006)	0.335 (0.006)	0.419 (0.006)	0.419 (0.006)
0.5	clust	0.761 (0.004)	0.780 (0.003)	0.304 (0.002)	0.730 (0.003)	0.407 (0.004)	0.534 (0.006)
0.5	strict-lassclust	0.761 (0.004)	0.781 (0.003)	0.302 (0.002)	0.730 (0.003)	0.407 (0.005)	0.533 (0.006)
0.5	non-strict-lassclust	0.761 (0.004)	0.781 (0.003)	0.302 (0.002)	0.730 (0.003)	0.407 (0.005)	0.533 (0.006)
0.5	lassojoint	0.804 (0.002)	0.804 (0.002)	0.794 (0.005)	0.794 (0.005)	0.504 (0.016)	0.504 (0.016)
0.5	spy	0.728 (0.015)		0.762 (0.005)		0.414 (0.136)	
0.8	naive	0.802 (0.001)	0.802 (0.001)	0.783 (0.001)	0.783 (0.001)	0.544 (0.001)	0.544 (0.001)
0.8	clust	0.755 (0.001)	0.789 (0.000)	0.241 (0.003)	0.785 (0.003)	0.383 (0.003)	0.493 (0.000)
0.8	strict-lassclust	0.755 (0.001)	0.789 (0.000)	0.241 (0.002)	0.786 (0.003)	0.383 (0.003)	0.493 (0.000)
0.8	non-strict-lassclust	0.755 (0.001)	0.789 (0.000)	0.241 (0.002)	0.786 (0.003)	0.383 (0.003)	0.493 (0.000)
0.8	lassojoint	0.801 (0.002)	0.801 (0.002)	0.796 (0.002)	0.796 (0.002)	0.470 (0.003)	0.470 (0.003)
0.8	spy	0.723 (0.030)		0.761 (0.009)		0.316 (0.152)	

Table 46. Summary Statistics for df = artif; SCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.968 (0.010)	0.975 (0.008)	0.520 (0.019)	0.821 (0.020)	0.670 (0.016)	0.844 (0.017)
0.3	clust	0.599 (0.063)	0.822 (0.060)	0.543 (0.037)	0.687 (0.037)	0.631 (0.035)	0.579 (0.069)
0.3	strict-lassclust	0.597 (0.065)	0.823 (0.060)	0.537 (0.043)	0.685 (0.055)	0.617 (0.057)	0.576 (0.139)
0.3	non-strict-lassclust	0.602 (0.063)	0.821 (0.058)	0.548 (0.039)	0.682 (0.041)	0.630 (0.036)	0.565 (0.087)
0.3	lassojoint	0.988 (0.011)	0.984 (0.011)	0.943 (0.016)	0.920 (0.078)	0.941 (0.019)	0.925 (0.050)
0.3	spy	0.829 (0.016)		0.712 (0.059)		0.666 (0.147)	
0.5	naive	0.983 (0.005)	0.985 (0.004)	0.627 (0.007)	0.808 (0.010)	0.729 (0.006)	0.839 (0.006)
0.5	clust	0.555 (0.017)	0.785 (0.024)	0.531 (0.002)	0.683 (0.032)	0.653 (0.003)	0.601 (0.045)
0.5	strict-lassclust	0.561 (0.016)	0.787 (0.025)	0.534 (0.007)	0.681 (0.040)	0.577 (0.057)	0.589 (0.099)
0.5	non-strict-lassclust	0.550 (0.018)	0.786 (0.022)	0.527 (0.004)	0.686 (0.031)	0.655 (0.012)	0.605 (0.045)
0.5	lassojoint	0.993 (0.001)	0.992 (0.001)	0.943 (0.003)	0.946 (0.002)	0.943 (0.003)	0.947 (0.002)
0.5	spy	0.831 (0.012)		0.738 (0.016)		0.722 (0.034)	
0.8	naive	0.988 (0.000)	0.988 (0.000)	0.883 (0.005)	0.883 (0.005)	0.890 (0.005)	0.890 (0.005)
0.8	clust	0.602 (0.012)	0.950 (0.025)	0.477 (0.003)	0.759 (0.008)	0.642 (0.002)	0.667 (0.007)
0.8	strict-lassclust	0.605 (0.010)	0.949 (0.025)	0.479 (0.004)	0.758 (0.024)	0.645 (0.001)	0.663 (0.047)
0.8	non-strict-lassclust	0.603 (0.013)	0.949 (0.024)	0.478 (0.002)	0.761 (0.008)	0.643 (0.002)	0.671 (0.002)
0.8	lassojoint	0.974 (0.000)	0.992 (0.005)	0.576 (0.005)	0.950 (0.070)	0.693 (0.004)	0.952 (0.046)
0.8	spy	0.831 (0.012)		0.742 (0.014)		0.724 (0.020)	

Table 47. Summary Statistics for df = bank-marketing; SCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.890 (0.023)	0.891 (0.021)	0.111 (0.006)	0.259 (0.109)	0.200 (0.010)	0.233 (0.035)
0.3	clust	0.718 (0.049)	0.835 (0.016)	0.623 (0.043)	0.753 (0.031)	0.305 (0.021)	0.380 (0.031)
0.3	strict-lassclust	0.718 (0.050)	0.838 (0.016)	0.629 (0.028)	0.753 (0.036)	0.307 (0.019)	0.381 (0.036)
0.3	non-strict-lassclust	0.717 (0.050)	0.838 (0.016)	0.628 (0.034)	0.750 (0.032)	0.307 (0.020)	0.378 (0.032)
0.3	lassojoint	0.896 (0.023)	0.891 (0.035)	0.798 (0.243)	0.861 (0.168)	0.436 (0.110)	0.448 (0.084)
0.3	spy	0.743 (0.024)		0.705 (0.013)		0.348 (0.021)	
0.5	naive	0.910 (0.003)	0.910 (0.003)	0.407 (0.077)	0.407 (0.077)	0.265 (0.028)	0.265 (0.028)
0.5	clust	0.739 (0.038)	0.841 (0.005)	0.585 (0.104)	0.745 (0.015)	0.283 (0.026)	0.363 (0.019)
0.5	strict-lassclust	0.740 (0.038)	0.843 (0.006)	0.586 (0.099)	0.745 (0.014)	0.282 (0.026)	0.364 (0.018)
0.5	non-strict-lassclust	0.739 (0.038)	0.843 (0.006)	0.586 (0.099)	0.745 (0.014)	0.282 (0.026)	0.364 (0.018)
0.5	lassojoint	0.906 (0.006)	0.906 (0.006)	0.902 (0.004)	0.902 (0.004)	0.509 (0.028)	0.510 (0.028)
0.5	spy	0.745 (0.023)		0.707 (0.013)		0.345 (0.023)	
0.8	naive	0.920 (0.004)	0.921 (0.004)	0.909 (0.006)	0.831 (0.010)	0.547 (0.027)	0.531 (0.019)
0.8	clust	0.680 (0.022)	0.862 (0.006)	0.124 (0.033)	0.868 (0.003)	0.195 (0.010)	0.514 (0.007)
0.8	strict-lassclust	0.674 (0.033)	0.864 (0.006)	0.133 (0.039)	0.868 (0.003)	0.196 (0.011)	0.514 (0.012)
0.8	non-strict-lassclust	0.669 (0.023)	0.864 (0.006)	0.116 (0.021)	0.868 (0.003)	0.194 (0.007)	0.514 (0.012)
0.8	lassojoint	0.921 (0.004)	0.922 (0.003)	0.915 (0.004)	0.904 (0.003)	0.511 (0.032)	0.578 (0.031)
0.8	spy	0.746 (0.021)		0.710 (0.013)		0.345 (0.023)	

Table 48. Summary Statistics for df = banknote; SCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.997 (0.002)	0.998 (0.002)	0.462 (0.019)	0.905 (0.019)	0.624 (0.018)	0.904 (0.018)
0.3	clust	0.680 (0.040)	0.786 (0.052)	0.586 (0.059)	0.750 (0.039)	0.578 (0.033)	0.683 (0.056)
0.3	strict-lassclust	0.677 (0.038)	0.785 (0.052)	0.580 (0.050)	0.749 (0.039)	0.575 (0.030)	0.681 (0.059)
0.3	non-strict-lassclust	0.677 (0.038)	0.786 (0.052)	0.580 (0.050)	0.749 (0.039)	0.574 (0.030)	0.682 (0.055)
0.3	lassojoint	0.999 (0.001)	0.999 (0.002)	0.985 (0.007)	0.984 (0.008)	0.983 (0.008)	0.982 (0.009)
0.3	spy	0.783 (0.096)		0.678 (0.062)		0.555 (0.108)	
0.5	naive	0.997 (0.001)	0.998 (0.001)	0.723 (0.013)	0.921 (0.011)	0.757 (0.009)	0.916 (0.010)
0.5	clust	0.663 (0.028)	0.794 (0.014)	0.565 (0.023)	0.770 (0.013)	0.575 (0.018)	0.690 (0.017)
0.5	strict-lassclust	0.660 (0.028)	0.794 (0.014)	0.561 (0.021)	0.769 (0.013)	0.575 (0.018)	0.689 (0.019)
0.5	non-strict-lassclust	0.660 (0.028)	0.794 (0.014)	0.561 (0.021)	0.769 (0.013)	0.575 (0.018)	0.689 (0.019)
0.5	lassojoint	0.999 (0.000)	0.999 (0.000)	0.982 (0.006)	0.981 (0.005)	0.980 (0.007)	0.979 (0.005)
0.5	spy	0.787 (0.086)		0.673 (0.055)		0.529 (0.099)	
0.8	naive	0.996 (0.001)	0.996 (0.001)	0.949 (0.000)	0.949 (0.000)	0.945 (0.000)	0.945 (0.000)
0.8	clust	0.687 (0.018)	0.863 (0.004)	0.522 (0.024)	0.815 (0.001)	0.611 (0.005)	0.753 (0.001)
0.8	strict-lassclust	0.681 (0.014)	0.862 (0.004)	0.529 (0.015)	0.815 (0.001)	0.618 (0.001)	0.753 (0.001)
0.8	non-strict-lassclust	0.681 (0.014)	0.862 (0.004)	0.529 (0.015)	0.815 (0.001)	0.618 (0.001)	0.753 (0.001)
0.8	lassojoint	1.000 (0.000)	1.000 (0.000)	0.981 (0.002)	0.988 (0.000)	0.978 (0.002)	0.986 (0.001)
0.8	spy	0.767 (0.093)		0.659 (0.051)		0.502 (0.094)	

Table 49. Summary Statistics for df = breastc; SCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.984 (0.011)	0.985 (0.011)	0.348 (0.031)	0.881 (0.037)	0.514 (0.034)	0.854 (0.040)
0.3	clust	0.979 (0.020)	0.983 (0.021)	0.964 (0.013)	0.956 (0.014)	0.947 (0.020)	0.934 (0.021)
0.3	strict-lassclust	0.994 (0.003)	0.994 (0.003)	0.966 (0.016)	0.958 (0.013)	0.951 (0.023)	0.937 (0.021)
0.3	non-strict-lassclust	0.994 (0.003)	0.994 (0.003)	0.966 (0.016)	0.958 (0.013)	0.951 (0.023)	0.937 (0.020)
0.3	lassojoint	0.983 (0.016)	0.983 (0.013)	0.868 (0.191)	0.857 (0.204)	0.867 (0.134)	0.860 (0.139)
0.3	spy	0.989 (0.006)		0.919 (0.023)		0.895 (0.029)	
0.5	naive	0.993 (0.003)	0.994 (0.002)	0.706 (0.032)	0.913 (0.041)	0.691 (0.024)	0.886 (0.037)
0.5	clust	0.988 (0.009)	0.997 (0.003)	0.963 (0.006)	0.968 (0.006)	0.946 (0.008)	0.951 (0.010)
0.5	strict-lassclust	0.995 (0.001)	0.997 (0.001)	0.965 (0.006)	0.970 (0.007)	0.949 (0.009)	0.954 (0.011)
0.5	non-strict-lassclust	0.995 (0.001)	0.997 (0.001)	0.965 (0.006)	0.970 (0.007)	0.949 (0.009)	0.954 (0.011)
0.5	lassojoint	0.996 (0.001)	0.996 (0.001)	0.958 (0.010)	0.961 (0.009)	0.940 (0.014)	0.943 (0.012)
0.5	spy	0.989 (0.006)		0.921 (0.026)		0.898 (0.031)	
0.8	naive	0.992 (0.003)	0.992 (0.003)	0.941 (0.015)	0.941 (0.015)	0.921 (0.016)	0.921 (0.016)
0.8	clust	0.962 (0.016)	0.992 (0.003)	0.940 (0.038)	0.960 (0.007)	0.919 (0.047)	0.944 (0.008)
0.8	strict-lassclust	0.988 (0.004)	0.991 (0.003)	0.867 (0.083)	0.962 (0.010)	0.841 (0.078)	0.947 (0.012)
0.8	non-strict-lassclust	0.988 (0.004)	0.991 (0.003)	0.867 (0.083)	0.962 (0.010)	0.841 (0.078)	0.947 (0.012)
0.8	lassojoint	0.993 (0.002)	0.993 (0.002)	0.934 (0.134)	0.934 (0.134)	0.929 (0.094)	0.929 (0.094)
0.8	spy	0.988 (0.006)		0.916 (0.022)		0.892 (0.029)	

Table 50. Summary Statistics for df = credit-a; SCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.813 (0.030)	0.818 (0.028)	0.457 (0.029)	0.740 (0.026)	0.626 (0.027)	0.730 (0.031)
0.3	clust	0.752 (0.048)	0.787 (0.037)	0.659 (0.029)	0.663 (0.034)	0.477 (0.064)	0.455 (0.066)
0.3	strict-lassclust	0.762 (0.033)	0.791 (0.030)	0.652 (0.031)	0.660 (0.034)	0.451 (0.065)	0.447 (0.065)
0.3	non-strict-lassclust	0.762 (0.033)	0.791 (0.030)	0.652 (0.031)	0.660 (0.034)	0.451 (0.065)	0.447 (0.065)
0.3	lassojoint	0.808 (0.032)	0.795 (0.040)	0.575 (0.139)	0.700 (0.096)	0.643 (0.048)	0.654 (0.051)
0.3	spy	0.769 (0.038)		0.703 (0.037)		0.607 (0.077)	
0.5	naive	0.800 (0.017)	0.801 (0.017)	0.450 (0.007)	0.707 (0.018)	0.618 (0.007)	0.697 (0.017)
0.5	clust	0.750 (0.006)	0.779 (0.013)	0.648 (0.012)	0.643 (0.006)	0.425 (0.057)	0.380 (0.037)
0.5	strict-lassclust	0.756 (0.007)	0.778 (0.013)	0.634 (0.009)	0.642 (0.007)	0.356 (0.047)	0.378 (0.042)
0.5	non-strict-lassclust	0.756 (0.007)	0.778 (0.013)	0.634 (0.009)	0.642 (0.007)	0.356 (0.047)	0.378 (0.042)
0.5	lassojoint	0.804 (0.015)	0.798 (0.016)	0.480 (0.092)	0.695 (0.038)	0.620 (0.008)	0.670 (0.016)
0.5	spy	0.774 (0.036)		0.709 (0.040)		0.603 (0.083)	
0.8	naive	0.837 (0.000)	0.837 (0.000)	0.764 (0.000)	0.764 (0.000)	0.744 (0.000)	0.744 (0.000)
0.8	clust	0.804 (0.000)	0.823 (0.000)	0.744 (0.000)	0.651 (0.000)	0.719 (0.000)	0.424 (0.000)
0.8	strict-lassclust	0.807 (0.000)	0.822 (0.000)	0.733 (0.000)	0.651 (0.000)	0.683 (0.000)	0.424 (0.000)
0.8	non-strict-lassclust	0.807 (0.000)	0.822 (0.000)	0.733 (0.000)	0.651 (0.000)	0.683 (0.000)	0.424 (0.000)
0.8	lassojoint	0.835 (0.000)	0.835 (0.000)	0.713 (0.000)	0.713 (0.000)	0.743 (0.000)	0.743 (0.000)
0.8	spy	0.774 (0.040)		0.713 (0.041)		0.615 (0.079)	

Table 51. Summary Statistics for df = credit-g; SCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.679 (0.042)	0.679 (0.041)	0.299 (0.023)	0.511 (0.057)	0.460 (0.027)	0.500 (0.034)
0.3	clust	0.619 (0.034)	0.642 (0.037)	0.631 (0.096)	0.708 (0.031)	0.414 (0.040)	0.307 (0.062)
0.3	strict-lassclust	0.617 (0.034)	0.633 (0.037)	0.637 (0.100)	0.707 (0.031)	0.409 (0.056)	0.291 (0.073)
0.3	non-strict-lassclust	0.619 (0.034)	0.639 (0.043)	0.636 (0.103)	0.707 (0.031)	0.403 (0.053)	0.291 (0.085)
0.3	lassojoint	0.652 (0.074)	0.656 (0.066)	0.596 (0.179)	0.696 (0.075)	0.436 (0.147)	0.298 (0.099)
0.3	spy	0.608 (0.038)		0.648 (0.044)		0.409 (0.047)	
0.5	naive	0.748 (0.011)	0.748 (0.011)	0.387 (0.029)	0.387 (0.029)	0.476 (0.017)	0.476 (0.017)
0.5	clust	0.639 (0.012)	0.663 (0.008)	0.636 (0.021)	0.741 (0.017)	0.457 (0.020)	0.407 (0.040)
0.5	strict-lassclust	0.643 (0.013)	0.662 (0.005)	0.653 (0.026)	0.743 (0.016)	0.451 (0.021)	0.410 (0.037)
0.5	non-strict-lassclust	0.644 (0.010)	0.661 (0.005)	0.656 (0.021)	0.743 (0.016)	0.452 (0.023)	0.410 (0.037)
0.5	lassojoint	0.741 (0.018)	0.741 (0.018)	0.349 (0.148)	0.349 (0.148)	0.424 (0.068)	0.424 (0.068)
0.5	spy	0.614 (0.034)		0.655 (0.034)		0.411 (0.039)	
0.8	naive	0.751 (0.022)	0.751 (0.022)	0.754 (0.017)	0.754 (0.017)	0.584 (0.002)	0.584 (0.002)
0.8	clust	0.592 (0.022)	0.656 (0.028)	0.366 (0.028)	0.722 (0.005)	0.454 (0.011)	0.290 (0.016)
0.8	strict-lassclust	0.623 (0.037)	0.651 (0.025)	0.454 (0.075)	0.715 (0.002)	0.469 (0.021)	0.247 (0.001)
0.8	non-strict-lassclust	0.611 (0.026)	0.647 (0.029)	0.418 (0.040)	0.718 (0.002)	0.464 (0.016)	0.262 (0.014)
0.8	lassojoint	0.735 (0.010)	0.735 (0.010)	0.727 (0.004)	0.727 (0.004)	0.220 (0.010)	0.220 (0.010)
0.8	spy	0.609 (0.032)		0.657 (0.025)		0.404 (0.037)	

Table 52. Summary Statistics for df = dhfr; SCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.592 (0.074)	0.604 (0.057)	0.534 (0.083)	0.547 (0.062)	0.559 (0.068)	0.561 (0.054)
0.3	clust	0.521 (0.117)	0.571 (0.064)	0.504 (0.164)	0.575 (0.071)	0.475 (0.036)	0.478 (0.071)
0.3	strict-lassclust	0.622 (0.072)	0.652 (0.108)	0.566 (0.092)	0.674 (0.059)	0.525 (0.122)	0.576 (0.101)
0.3	non-strict-lassclust	0.614 (0.077)	0.652 (0.098)	0.534 (0.094)	0.618 (0.131)	0.486 (0.081)	0.501 (0.099)
0.3	lassojoint	0.837 (0.095)	0.876 (0.050)	0.643 (0.175)	0.549 (0.119)	0.646 (0.122)	0.608 (0.056)
0.3	spy	0.664 (0.049)		0.628 (0.067)		0.498 (0.064)	
0.5	naive	0.647 (0.012)	0.624 (0.032)	0.596 (0.014)	0.585 (0.034)	0.631 (0.003)	0.602 (0.018)
0.5	clust	0.476 (0.009)	0.647 (0.004)	0.443 (0.023)	0.654 (0.019)	0.465 (0.038)	0.567 (0.007)
0.5	strict-lassclust	0.685 (0.022)	0.750 (0.014)	0.460 (0.013)	0.767 (0.010)	0.546 (0.028)	0.655 (0.020)
0.5	non-strict-lassclust	0.685 (0.032)	0.733 (0.031)	0.527 (0.056)	0.725 (0.034)	0.516 (0.004)	0.618 (0.034)
0.5	lassojoint	0.954 (0.027)	0.942 (0.006)	0.783 (0.048)	0.750 (0.060)	0.776 (0.027)	0.749 (0.051)
0.5	spy	0.679 (0.075)		0.645 (0.098)		0.545 (0.089)	
0.8	naive	0.652 (0.000)	0.684 (0.000)	0.619 (0.000)	0.660 (0.000)	0.584 (0.000)	0.612 (0.000)
0.8	clust	0.481 (0.000)	0.600 (0.000)	0.443 (0.000)	0.649 (0.000)	0.426 (0.000)	0.514 (0.000)
0.8	strict-lassclust	0.658 (0.000)	0.746 (0.000)	0.412 (0.000)	0.784 (0.000)	0.513 (0.000)	0.644 (0.000)
0.8	non-strict-lassclust	0.658 (0.000)	0.748 (0.000)	0.423 (0.000)	0.784 (0.000)	0.517 (0.000)	0.644 (0.000)
0.8	lassojoint	0.929 (0.000)	0.943 (0.000)	0.567 (0.000)	0.887 (0.000)	0.611 (0.000)	0.836 (0.000)
0.8	spy	0.720 (0.039)		0.642 (0.049)		0.600 (0.045)	

Table 53. Summary Statistics for df = diabetes; SCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.794 (0.032)	0.800 (0.031)	0.352 (0.030)	0.642 (0.042)	0.516 (0.032)	0.629 (0.038)
0.3	clust	0.727 (0.049)	0.769 (0.032)	0.626 (0.055)	0.714 (0.028)	0.595 (0.049)	0.504 (0.069)
0.3	strict-lassclust	0.725 (0.050)	0.770 (0.034)	0.634 (0.058)	0.709 (0.033)	0.588 (0.055)	0.470 (0.127)
0.3	non-strict-lassclust	0.731 (0.049)	0.768 (0.033)	0.610 (0.070)	0.711 (0.030)	0.591 (0.047)	0.505 (0.071)
0.3	lassojoint	0.800 (0.033)	0.789 (0.043)	0.684 (0.129)	0.743 (0.049)	0.595 (0.062)	0.570 (0.071)
0.3	spy	0.677 (0.062)		0.648 (0.051)		0.534 (0.089)	
0.5	naive	0.791 (0.014)	0.791 (0.014)	0.533 (0.037)	0.546 (0.050)	0.588 (0.025)	0.591 (0.027)
0.5	clust	0.723 (0.022)	0.757 (0.020)	0.551 (0.067)	0.687 (0.013)	0.594 (0.033)	0.495 (0.047)
0.5	strict-lassclust	0.723 (0.021)	0.755 (0.020)	0.539 (0.066)	0.688 (0.016)	0.590 (0.033)	0.487 (0.053)
0.5	non-strict-lassclust	0.725 (0.020)	0.756 (0.020)	0.535 (0.068)	0.694 (0.015)	0.588 (0.033)	0.512 (0.054)
0.5	lassojoint	0.791 (0.015)	0.793 (0.013)	0.707 (0.041)	0.707 (0.040)	0.639 (0.012)	0.638 (0.016)
0.5	spy	0.690 (0.054)		0.662 (0.049)		0.546 (0.080)	
0.8	naive	0.797 (0.019)	0.797 (0.019)	0.724 (0.016)	0.724 (0.016)	0.653 (0.016)	0.653 (0.016)
0.8	clust	0.712 (0.037)	0.767 (0.024)	0.346 (0.008)	0.678 (0.030)	0.507 (0.014)	0.390 (0.055)
0.8	strict-lassclust	0.703 (0.042)	0.766 (0.025)	0.348 (0.005)	0.671 (0.034)	0.510 (0.015)	0.366 (0.073)
0.8	non-strict-lassclust	0.707 (0.047)	0.766 (0.025)	0.344 (0.011)	0.671 (0.033)	0.509 (0.016)	0.366 (0.072)
0.8	lassojoint	0.798 (0.015)	0.798 (0.015)	0.352 (0.016)	0.353 (0.016)	0.512 (0.016)	0.512 (0.016)
0.8	spy	0.678 (0.049)		0.661 (0.038)		0.527 (0.065)	

Table 54. Summary Statistics for df = heart-c; SCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.708 (0.065)	0.709 (0.060)	0.466 (0.040)	0.631 (0.053)	0.633 (0.038)	0.655 (0.051)
0.3	clust	0.710 (0.055)	0.722 (0.056)	0.603 (0.062)	0.647 (0.056)	0.674 (0.058)	0.547 (0.107)
0.3	strict-lassclust	0.708 (0.055)	0.722 (0.056)	0.600 (0.060)	0.646 (0.056)	0.673 (0.058)	0.543 (0.106)
0.3	non-strict-lassclust	0.708 (0.055)	0.722 (0.056)	0.600 (0.060)	0.646 (0.056)	0.673 (0.058)	0.543 (0.106)
0.3	lassojoint	0.677 (0.102)	0.685 (0.071)	0.563 (0.096)	0.611 (0.071)	0.608 (0.112)	0.491 (0.136)
0.3	spy	0.694 (0.075)		0.606 (0.065)		0.427 (0.199)	
0.5	naive	0.739 (0.003)	0.733 (0.001)	0.433 (0.003)	0.578 (0.003)	0.605 (0.003)	0.655 (0.003)
0.5	clust	0.731 (0.005)	0.749 (0.005)	0.456 (0.000)	0.700 (0.002)	0.608 (0.002)	0.629 (0.010)
0.5	strict-lassclust	0.731 (0.005)	0.749 (0.005)	0.444 (0.000)	0.700 (0.002)	0.603 (0.002)	0.629 (0.010)
0.5	non-strict-lassclust	0.731 (0.005)	0.749 (0.005)	0.444 (0.000)	0.700 (0.002)	0.603 (0.002)	0.629 (0.010)
0.5	lassojoint	0.727 (0.006)	0.730 (0.006)	0.686 (0.029)	0.690 (0.007)	0.631 (0.006)	0.549 (0.014)
0.5	spy	0.695 (0.055)		0.600 (0.051)		0.409 (0.168)	
0.8	naive	0.746 (0.000)	0.746 (0.000)	0.611 (0.002)	0.611 (0.002)	0.599 (0.007)	0.599 (0.007)
0.8	clust	0.792 (0.005)	0.794 (0.004)	0.400 (0.003)	0.778 (0.005)	0.571 (0.003)	0.677 (0.001)
0.8	strict-lassclust	0.793 (0.005)	0.793 (0.004)	0.400 (0.003)	0.767 (0.003)	0.571 (0.003)	0.656 (0.004)
0.8	non-strict-lassclust	0.793 (0.005)	0.793 (0.004)	0.400 (0.003)	0.767 (0.003)	0.571 (0.003)	0.656 (0.004)
0.8	lassojoint	0.777 (0.003)	0.777 (0.003)	0.777 (0.011)	0.777 (0.011)	0.696 (0.014)	0.696 (0.014)
0.8	spy	0.705 (0.055)		0.602 (0.044)		0.372 (0.142)	

Table 55. Summary Statistics for df = qsar-biodeg; SCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.815 (0.031)	0.805 (0.035)	0.372 (0.025)	0.637 (0.042)	0.516 (0.024)	0.627 (0.031)
0.3	clust	0.778 (0.028)	0.836 (0.029)	0.546 (0.044)	0.767 (0.030)	0.583 (0.030)	0.678 (0.039)
0.3	strict-lassclust	0.781 (0.030)	0.841 (0.028)	0.543 (0.090)	0.748 (0.055)	0.551 (0.109)	0.602 (0.184)
0.3	non-strict-lassclust	0.773 (0.039)	0.838 (0.032)	0.506 (0.076)	0.733 (0.092)	0.568 (0.038)	0.663 (0.051)
0.3	lassojoint	0.841 (0.052)	0.820 (0.038)	0.746 (0.104)	0.773 (0.033)	0.674 (0.058)	0.661 (0.051)
0.3	spy	0.767 (0.076)		0.668 (0.098)		0.572 (0.161)	
0.5	naive	0.866 (0.008)	0.863 (0.011)	0.512 (0.026)	0.643 (0.098)	0.583 (0.015)	0.654 (0.054)
0.5	clust	0.798 (0.020)	0.863 (0.009)	0.582 (0.021)	0.790 (0.014)	0.617 (0.016)	0.707 (0.005)
0.5	strict-lassclust	0.818 (0.014)	0.873 (0.011)	0.626 (0.058)	0.720 (0.073)	0.633 (0.033)	0.476 (0.354)
0.5	non-strict-lassclust	0.772 (0.046)	0.859 (0.010)	0.466 (0.088)	0.672 (0.127)	0.566 (0.034)	0.669 (0.068)
0.5	lassojoint	0.877 (0.013)	0.856 (0.025)	0.832 (0.013)	0.811 (0.020)	0.753 (0.025)	0.693 (0.037)
0.5	spy	0.789 (0.058)		0.701 (0.053)		0.628 (0.075)	
0.8	naive	0.904 (0.000)	0.904 (0.000)	0.807 (0.000)	0.807 (0.000)	0.775 (0.000)	0.775 (0.000)
0.8	clust	0.807 (0.000)	0.916 (0.000)	0.522 (0.000)	0.839 (0.000)	0.612 (0.000)	0.783 (0.000)
0.8	strict-lassclust	0.806 (0.000)	0.911 (0.000)	0.443 (0.000)	0.816 (0.000)	0.577 (0.000)	0.743 (0.000)
0.8	non-strict-lassclust	0.806 (0.000)	0.911 (0.000)	0.440 (0.000)	0.816 (0.000)	0.576 (0.000)	0.743 (0.000)
0.8	lassojoint	0.916 (0.000)	0.916 (0.000)	0.851 (0.000)	0.851 (0.000)	0.795 (0.000)	0.795 (0.000)
0.8	spy	0.812 (0.047)		0.718 (0.035)		0.656 (0.038)	

Table 56. Summary Statistics for df = spambase; SCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.936 (0.010)	0.933 (0.010)	0.397 (0.011)	0.832 (0.020)	0.565 (0.011)	0.812 (0.019)
0.3	clust	0.896 (0.068)	0.921 (0.049)	0.803 (0.148)	0.857 (0.026)	0.787 (0.085)	0.797 (0.048)
0.3	strict-lassclust	0.906 (0.028)	0.937 (0.009)	0.792 (0.158)	0.857 (0.024)	0.783 (0.089)	0.798 (0.044)
0.3	non-strict-lassclust	0.912 (0.020)	0.937 (0.009)	0.794 (0.156)	0.859 (0.026)	0.783 (0.088)	0.803 (0.046)
0.3	lassojoint	0.946 (0.012)	0.943 (0.013)	0.863 (0.119)	0.892 (0.018)	0.846 (0.074)	0.862 (0.022)
0.3	spy	0.841 (0.029)		0.748 (0.067)		0.732 (0.045)	
0.5	naive	0.956 (0.004)	0.953 (0.005)	0.530 (0.025)	0.859 (0.010)	0.631 (0.012)	0.844 (0.010)
0.5	clust	0.918 (0.036)	0.855 (0.078)	0.864 (0.021)	0.825 (0.014)	0.828 (0.036)	0.737 (0.025)
0.5	strict-lassclust	0.916 (0.005)	0.941 (0.005)	0.855 (0.027)	0.824 (0.014)	0.827 (0.022)	0.734 (0.025)
0.5	non-strict-lassclust	0.916 (0.006)	0.941 (0.005)	0.856 (0.028)	0.824 (0.013)	0.827 (0.022)	0.733 (0.023)
0.5	lassojoint	0.968 (0.004)	0.958 (0.011)	0.918 (0.006)	0.855 (0.137)	0.897 (0.008)	0.847 (0.090)
0.5	spy	0.834 (0.025)		0.746 (0.065)		0.732 (0.042)	
0.8	naive	0.963 (0.001)	0.963 (0.001)	0.913 (0.003)	0.913 (0.003)	0.894 (0.003)	0.894 (0.003)
0.8	clust	0.851 (0.053)	0.942 (0.005)	0.724 (0.088)	0.842 (0.016)	0.737 (0.049)	0.769 (0.028)
0.8	strict-lassclust	0.907 (0.018)	0.941 (0.004)	0.619 (0.047)	0.838 (0.017)	0.671 (0.021)	0.760 (0.031)
0.8	non-strict-lassclust	0.910 (0.004)	0.941 (0.004)	0.620 (0.048)	0.838 (0.017)	0.672 (0.022)	0.761 (0.031)
0.8	lassojoint	0.942 (0.004)	0.942 (0.004)	0.402 (0.003)	0.402 (0.003)	0.572 (0.003)	0.572 (0.003)
0.8	spy	0.829 (0.020)		0.769 (0.037)		0.743 (0.026)	

Table 57. Summary Statistics for df = wdbc; SCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.959 (0.022)	0.958 (0.021)	0.435 (0.037)	0.748 (0.046)	0.566 (0.029)	0.742 (0.044)
0.3	clust	0.899 (0.028)	0.914 (0.021)	0.886 (0.028)	0.903 (0.021)	0.840 (0.039)	0.861 (0.033)
0.3	strict-lassclust	0.964 (0.011)	0.970 (0.013)	0.903 (0.026)	0.904 (0.028)	0.861 (0.046)	0.858 (0.051)
0.3	non-strict-lassclust	0.965 (0.012)	0.972 (0.011)	0.908 (0.018)	0.911 (0.020)	0.872 (0.025)	0.872 (0.031)
0.3	lassojoint	0.974 (0.021)	0.964 (0.025)	0.879 (0.127)	0.621 (0.239)	0.867 (0.092)	0.687 (0.149)
0.3	spy	0.959 (0.010)		0.900 (0.018)		0.869 (0.023)	
0.5	naive	0.970 (0.006)	0.967 (0.006)	0.684 (0.049)	0.839 (0.048)	0.731 (0.044)	0.838 (0.048)
0.5	clust	0.911 (0.004)	0.917 (0.010)	0.889 (0.007)	0.919 (0.005)	0.874 (0.016)	0.903 (0.007)
0.5	strict-lassclust	0.957 (0.004)	0.971 (0.007)	0.900 (0.020)	0.895 (0.017)	0.885 (0.028)	0.868 (0.029)
0.5	non-strict-lassclust	0.957 (0.004)	0.973 (0.005)	0.902 (0.020)	0.906 (0.016)	0.890 (0.022)	0.887 (0.016)
0.5	lassojoint	0.977 (0.003)	0.979 (0.009)	0.926 (0.015)	0.505 (0.099)	0.915 (0.021)	0.638 (0.044)
0.5	spy	0.958 (0.012)		0.899 (0.019)		0.866 (0.026)	
0.8	naive	0.983 (0.000)	0.983 (0.000)	0.912 (0.000)	0.912 (0.000)	0.887 (0.000)	0.887 (0.000)
0.8	clust	0.949 (0.000)	0.935 (0.000)	0.912 (0.000)	0.906 (0.000)	0.885 (0.000)	0.864 (0.000)
0.8	strict-lassclust	0.980 (0.000)	0.985 (0.000)	0.912 (0.000)	0.924 (0.000)	0.891 (0.000)	0.891 (0.000)
0.8	non-strict-lassclust	0.980 (0.000)	0.987 (0.000)	0.906 (0.000)	0.929 (0.000)	0.884 (0.000)	0.898 (0.000)
0.8	lassojoint	0.981 (0.000)	0.981 (0.000)	0.959 (0.000)	0.959 (0.000)	0.945 (0.000)	0.945 (0.000)
0.8	spy	0.959 (0.012)		0.902 (0.020)		0.872 (0.027)	

Table 58. Summary Statistics for df = wine-quality; SCAR scheme

calc-c	method	AUC		ACC		F1	
		nonSMOTE	SMOTE	nonSMOTE	SMOTE	nonSMOTE	SMOTE
0.3	naive	0.814 (0.046)	0.814 (0.042)	0.137 (0.013)	0.370 (0.125)	0.238 (0.019)	0.298 (0.044)
0.3	clust	0.694 (0.034)	0.737 (0.035)	0.445 (0.084)	0.777 (0.053)	0.280 (0.050)	0.369 (0.085)
0.3	strict-lassclust	0.695 (0.035)	0.736 (0.035)	0.460 (0.086)	0.781 (0.052)	0.283 (0.054)	0.363 (0.079)
0.3	non-strict-lassclust	0.696 (0.033)	0.735 (0.035)	0.458 (0.082)	0.777 (0.055)	0.283 (0.052)	0.361 (0.084)
0.3	lassojoint	0.807 (0.081)	0.801 (0.063)	0.710 (0.275)	0.818 (0.143)	0.376 (0.147)	0.345 (0.113)
0.3	spy	0.645 (0.059)		0.659 (0.122)		0.282 (0.070)	
0.5	naive	0.829 (0.013)	0.829 (0.013)	0.448 (0.034)	0.448 (0.034)	0.320 (0.019)	0.320 (0.019)
0.5	clust	0.701 (0.024)	0.759 (0.016)	0.279 (0.073)	0.801 (0.011)	0.257 (0.017)	0.403 (0.019)
0.5	strict-lassclust	0.707 (0.017)	0.752 (0.020)	0.290 (0.070)	0.803 (0.010)	0.264 (0.013)	0.393 (0.025)
0.5	non-strict-lassclust	0.701 (0.023)	0.752 (0.020)	0.295 (0.067)	0.804 (0.011)	0.260 (0.016)	0.394 (0.027)
0.5	lassojoint	0.824 (0.015)	0.820 (0.019)	0.818 (0.080)	0.834 (0.040)	0.398 (0.057)	0.392 (0.048)
0.5	spy	0.633 (0.062)		0.645 (0.129)		0.274 (0.066)	
0.8	naive	0.825 (0.003)	0.822 (0.001)	0.833 (0.000)	0.676 (0.007)	0.414 (0.007)	0.416 (0.007)
0.8	clust	0.674 (0.007)	0.772 (0.003)	0.161 (0.001)	0.801 (0.002)	0.254 (0.004)	0.444 (0.001)
0.8	strict-lassclust	0.674 (0.004)	0.773 (0.002)	0.149 (0.003)	0.800 (0.001)	0.251 (0.004)	0.435 (0.001)
0.8	non-strict-lassclust	0.673 (0.006)	0.773 (0.003)	0.148 (0.001)	0.799 (0.003)	0.251 (0.004)	0.436 (0.005)
0.8	lassojoint	0.830 (0.002)	0.815 (0.003)	0.852 (0.005)	0.818 (0.007)	0.381 (0.017)	0.435 (0.020)
0.8	spy	0.643 (0.058)		0.657 (0.099)		0.299 (0.066)	

2.3 Graphs

In this section we present all barplots for classification metrics for our methods. In addition, we present boxplots of the execution times of the algorithms used.

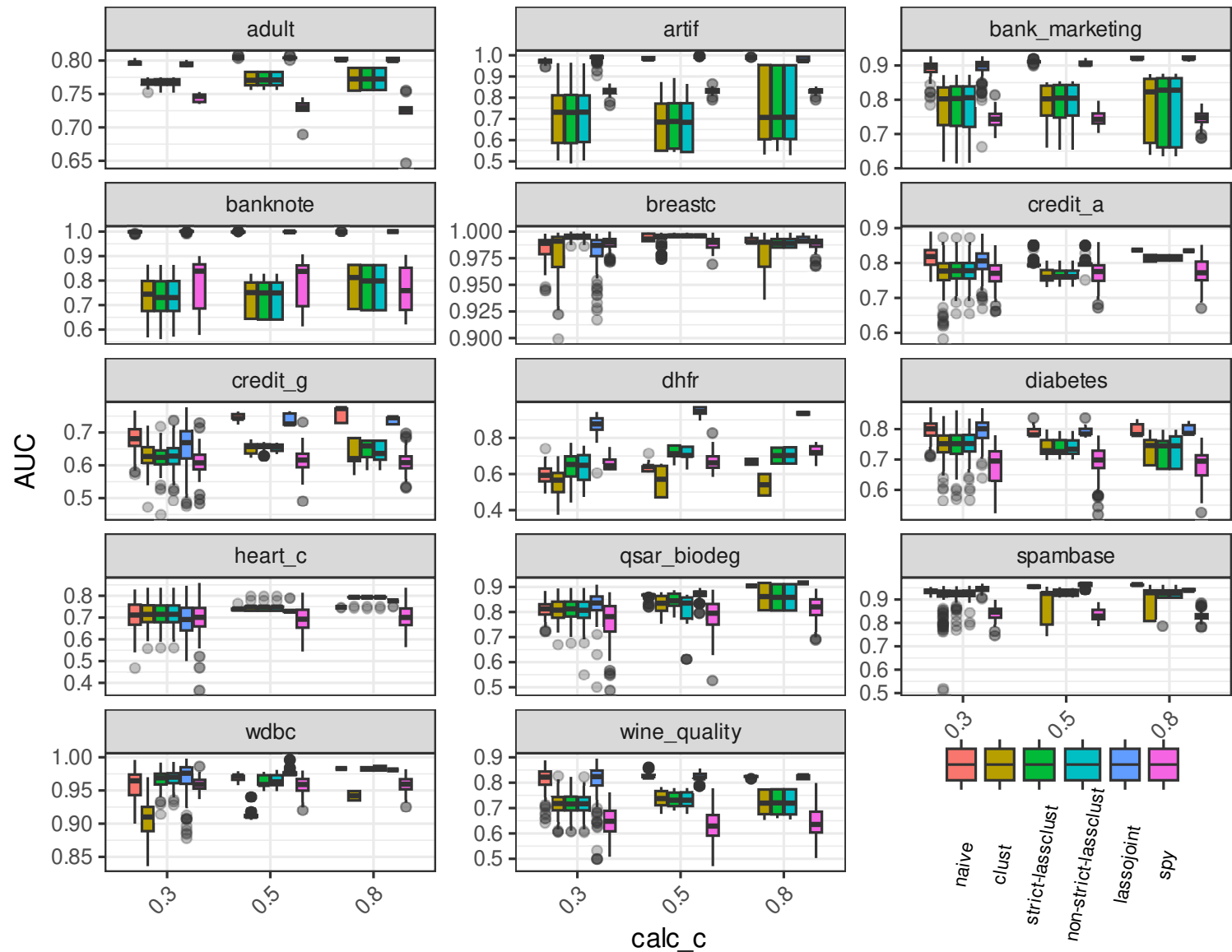


Fig. 4. AUC; SMOTE scheme; SCAR

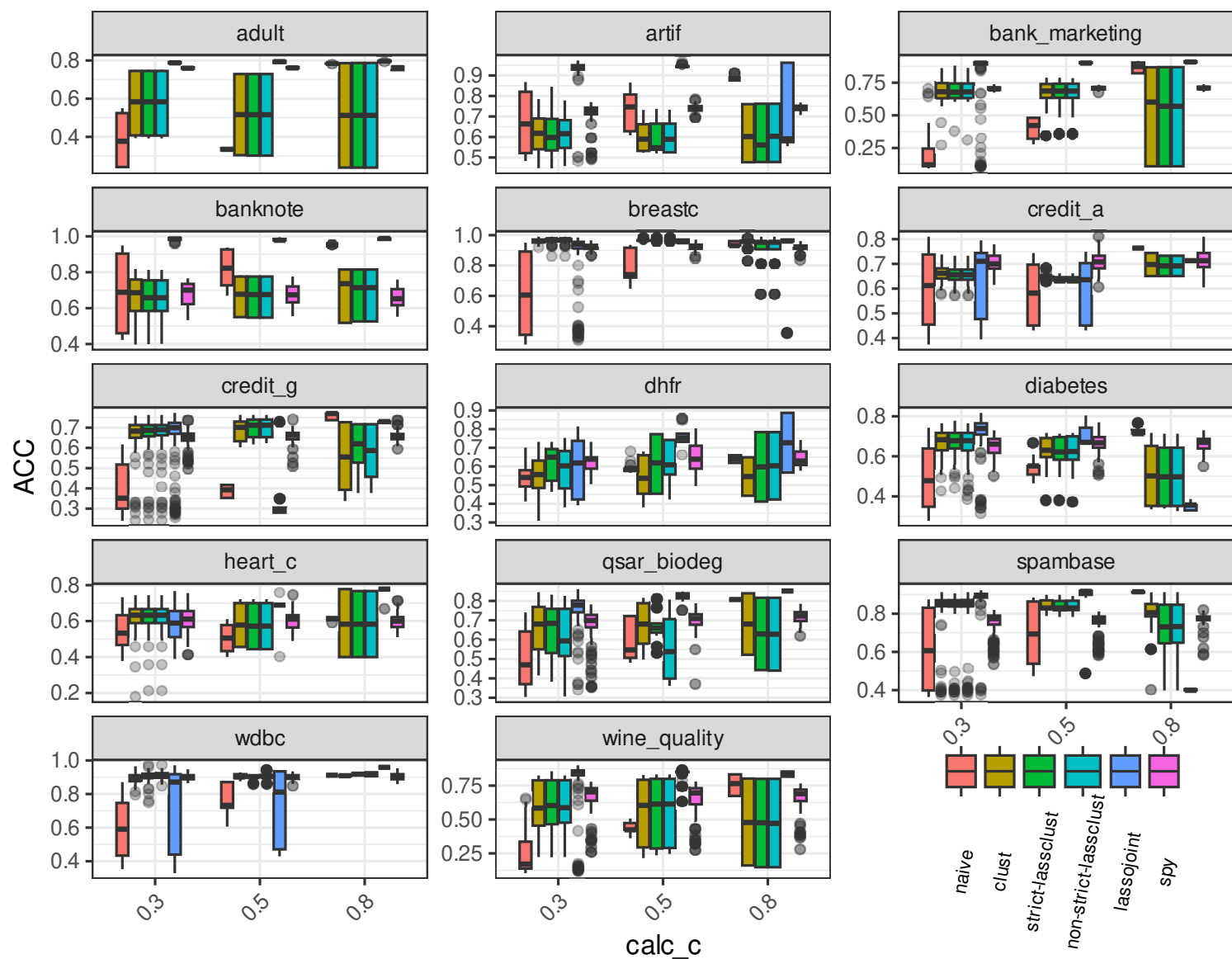


Fig. 5. Accuracy; SMOTE scheme; SCAR

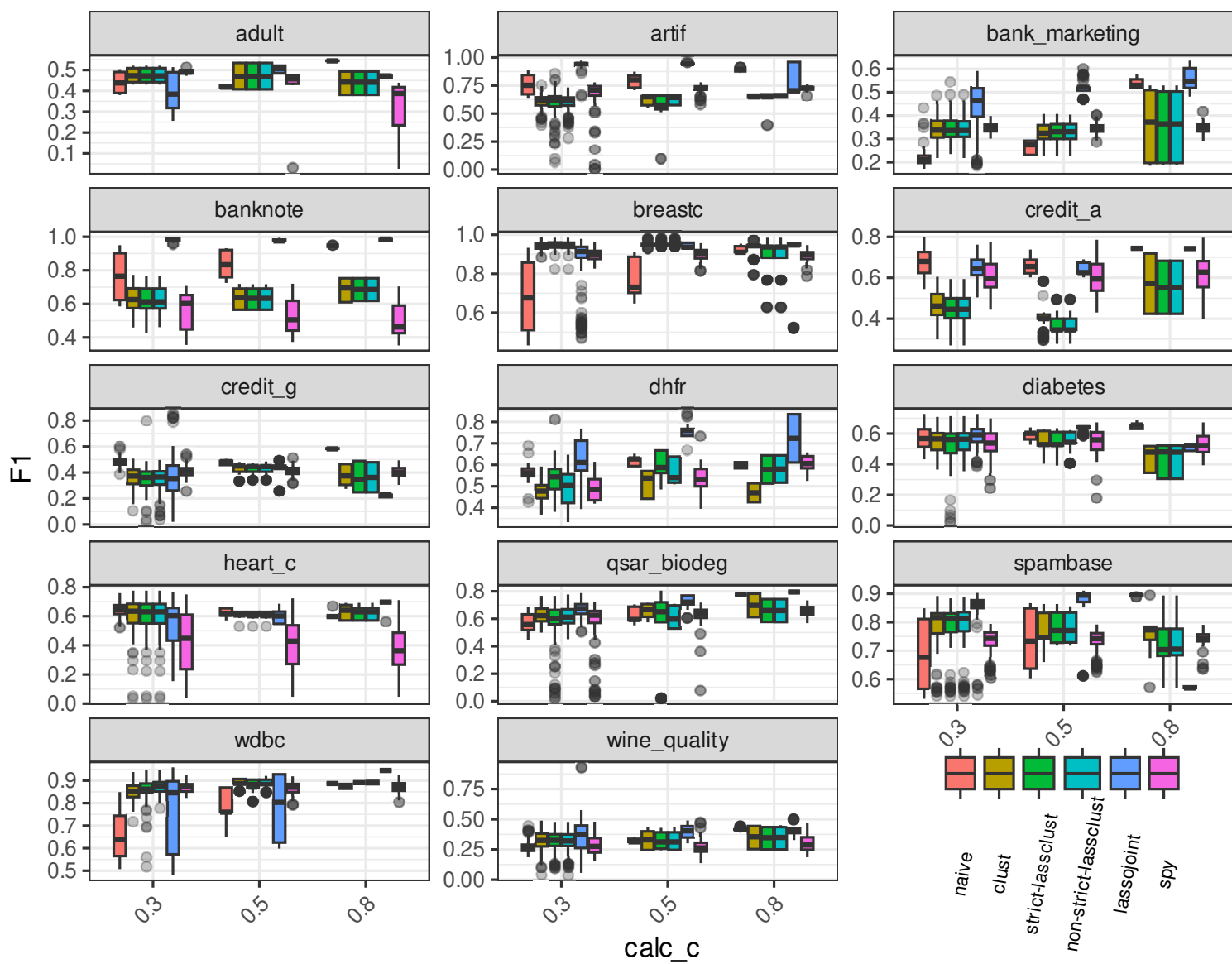


Fig. 6. F1 score; SMOTE scheme; SCAR

Calculation times

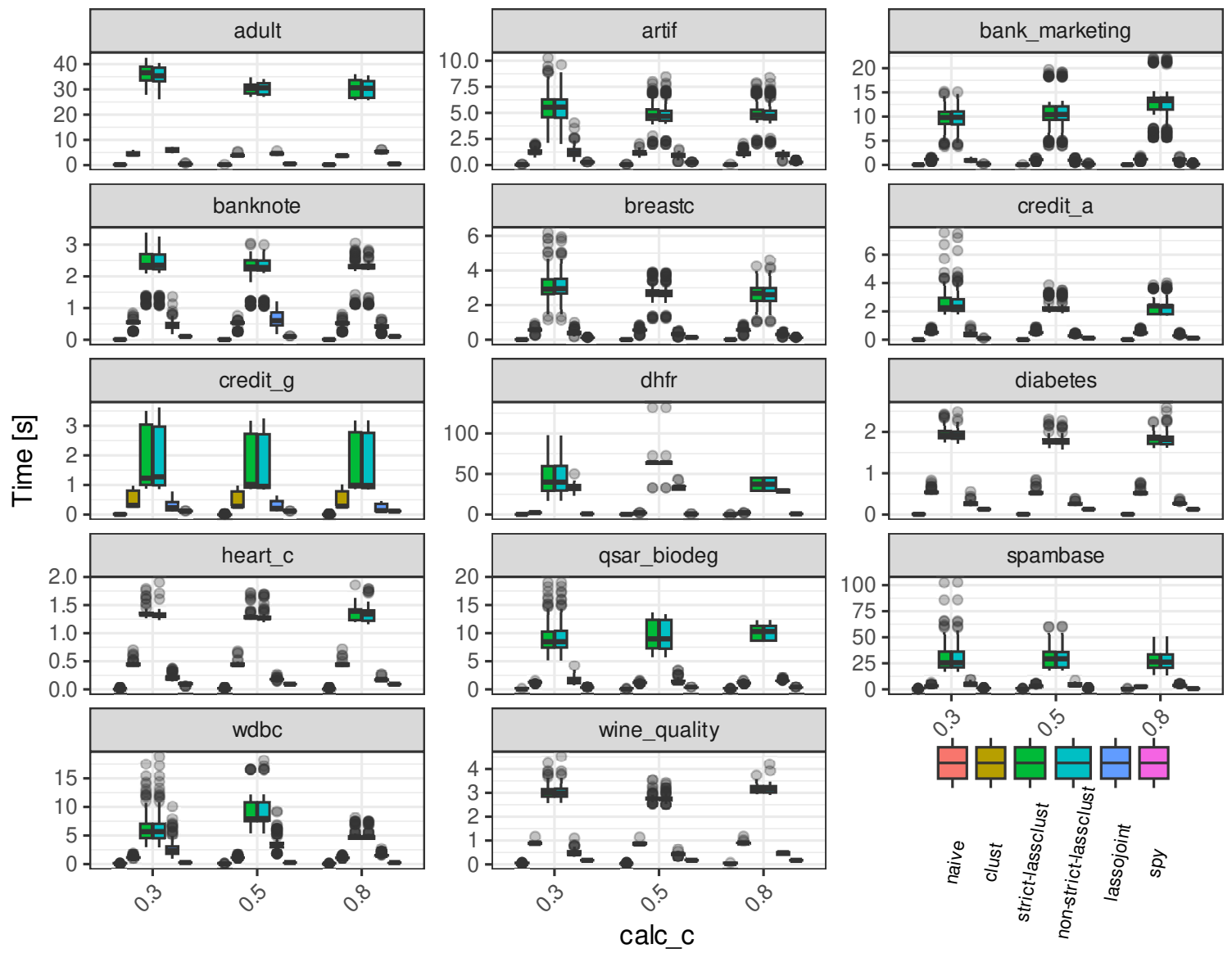


Fig. 7. Time[s]; SMOTE scheme; SCAR

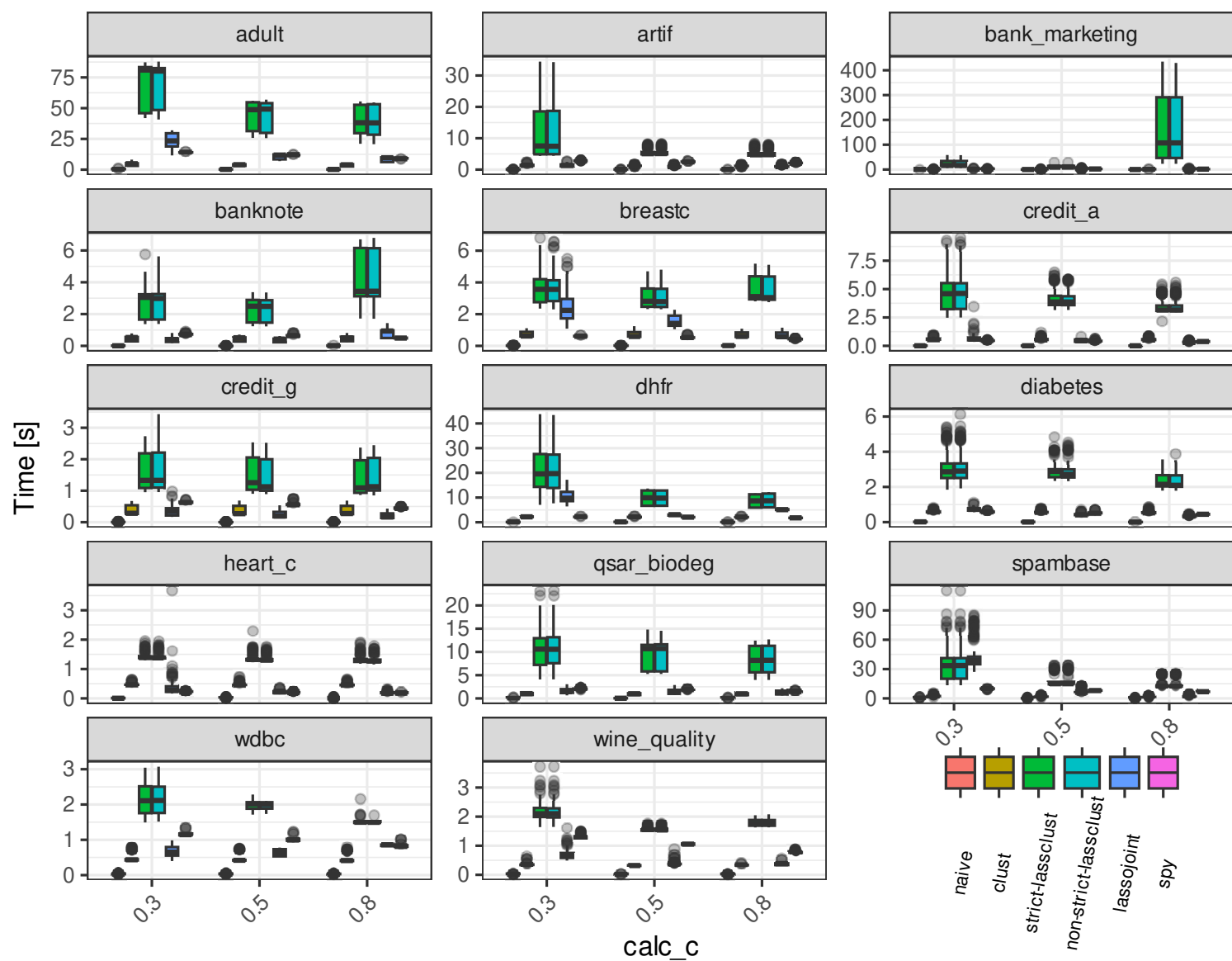


Fig. 8. Time[s]; SMOTE scheme; nonSCAR

4 Description of the Function `non_scar_labelling.mvc`

Given a dataset \mathbf{D} with columns representing features and a target variable Y , the function `non_scar_labelling.mvc` performs the following steps:

1. Filter out columns Y :

$$\mathbf{D}_{\text{filtered}} = \mathbf{D} \setminus \{Y\}$$

2. Calculate variance for each column:

$$\text{var}_i = \text{Var}(X_i) \quad \text{for each feature } X_i \in \mathbf{D}_{\text{filtered}}$$

3. Sort columns by decreasing variance:

$$\text{sorted_variances} = \text{sort}(\{\text{var}_i\}, \text{decreasing} = \text{TRUE})$$

4. Select top n_{vars} columns with highest variance:

$$\mathbf{T} = \{X_i \mid \text{var}_i \in \text{sorted_variances}[1 : n_{\text{vars}}]\}$$

5. Create temporary dataframe with new column defined as rank of sums of selected column, where rank $\text{rn}=1$ for the smallest value and rank $\text{rn}=n$ for the largest value and n is the number of rows in our dataset :

$$\mathbf{D}_{\text{temp}} = \mathbf{D} \cup \left\{ \text{rn} = \text{row_number} \left(\sum_{X_i \in \mathbf{T}} X_i \right) \right\}$$

6. Define binary variables S_{02} , S_{05} , and S_{08} :

$$S_{02} = \begin{cases} 1 & \text{if } \frac{\text{rn}}{n} < 0.2 \text{ and } Y = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$S_{04} = \begin{cases} 1 & \text{if } \frac{\text{rn}}{n} < 0.4 \text{ and } Y = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$S_{06} = \begin{cases} 1 & \text{if } \frac{\text{rn}}{n} < 0.6 \text{ and } Y = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$S_{08} = \begin{cases} 1 & \text{if } \frac{\text{rn}}{n} < 0.8 \text{ and } Y = 1 \\ 0 & \text{otherwise} \end{cases}$$

7. Calculate c_{calc} values:

$$c_{\text{calc}02} = \frac{\sum S_{02}}{\sum Y}, \quad c_{\text{calc}04} = \frac{\sum S_{04}}{\sum Y}, \quad c_{\text{calc}06} = \frac{\sum S_{06}}{\sum Y}, \quad c_{\text{calc}08} = \frac{\sum S_{08}}{\sum Y}$$

$$\mathbf{c}_{\text{calc}} = \{c_{\text{calc}02}, c_{\text{calc}04}, c_{\text{calc}06}, c_{\text{calc}08}\}$$

8. Fit a linear model to estimate target c_{calc} :

$$c_{\text{calc}} \sim \alpha + \beta \times x, \quad \text{where } x \in \{0.2, 0.4, 0.6, 0.8\}$$

$$\text{rn_frac} = \frac{\text{target_c_calc} - \alpha}{\beta}$$

9. Create the final dataframe with the new column S :

$$S = \begin{cases} 1 & \text{if } \frac{\text{rn}}{n} < \text{rn_frac} \text{ and } Y = 1 \\ 0 & \text{otherwise} \end{cases}$$

The final dataframe $\mathbf{D1} = \mathbf{D} \cup S$ is returned with the new column S representing the target variable's adjusted binary label.