

Table 1: The prediction performance under BCS.

Datasets		Higgs		Coverttype		Temp		Gas	
Models		Accuracy \uparrow	AUC \uparrow	Accuracy \uparrow	AUC \uparrow	RMSE \downarrow	R2 \uparrow	RMSE \downarrow	R2 \uparrow
F-XGBoost	F-XGBoost	0.649 \pm 0.003	0.707 \pm 0.002	0.707 \pm 0.005	0.933 \pm 0.003	1.281 \pm 0.016	0.730 \pm 0.004	103.436 \pm 2.551	0.994 \pm 0.001
	F-Mean	0.635 \pm 0.003	0.685 \pm 0.001	0.692 \pm 0.003	0.911 \pm 0.001	1.397 \pm 0.091	0.681 \pm 0.023	116.968 \pm 3.019	0.993 \pm 0.001
	F-MIWAE	0.618 \pm 0.004	0.666 \pm 0.005	0.678 \pm 0.003	0.897 \pm 0.004	1.629 \pm 0.073	0.567 \pm 0.016	—	—
	F-NMIWAE	0.621 \pm 0.005	0.669 \pm 0.003	0.683 \pm 0.001	0.901 \pm 0.002	1.598 \pm 0.011	0.583 \pm 0.019	—	—
F-RF	F-GAIN	0.638 \pm 0.000	0.689 \pm 0.003	0.697 \pm 0.004	0.919 \pm 0.003	1.347 \pm 0.031	0.687 \pm 0.001	107.149 \pm 5.147	0.994 \pm 0.001
	F-Mean	0.619 \pm 0.004	0.659 \pm 0.007	0.664 \pm 0.002	0.885 \pm 0.003	—	—	—	—
	F-MIWAE	0.608 \pm 0.004	0.646 \pm 0.005	0.655 \pm 0.001	0.869 \pm 0.003	—	—	—	—
	F-NMIWAE	0.610 \pm 0.001	0.648 \pm 0.002	0.659 \pm 0.003	0.874 \pm 0.001	—	—	—	—
F-MLP	F-GAIN	0.624 \pm 0.003	0.665 \pm 0.004	0.671 \pm 0.001	0.894 \pm 0.001	—	—	—	—
	F-Mean	0.623 \pm 0.004	0.666 \pm 0.003	0.677 \pm 0.003	0.894 \pm 0.001	1.560 \pm 0.019	0.603 \pm 0.008	133.398 \pm 7.269	0.990 \pm 0.001
	F-MIWAE	0.617 \pm 0.004	0.656 \pm 0.002	0.659 \pm 0.007	0.882 \pm 0.002	1.695 \pm 0.114	0.531 \pm 0.013	—	—
	F-NMIWAE	0.621 \pm 0.001	0.660 \pm 0.002	0.668 \pm 0.005	0.890 \pm 0.003	1.679 \pm 0.102	0.540 \pm 0.010	—	—
F-TabNet	F-GAIN	0.633 \pm 0.005	0.684 \pm 0.005	0.683 \pm 0.002	0.899 \pm 0.002	1.496 \pm 0.092	0.635 \pm 0.008	125.699 \pm 4.937	0.992 \pm 0.001
	F-Mean	0.629 \pm 0.003	0.674 \pm 0.003	0.717 \pm 0.002	0.937 \pm 0.001	1.391 \pm 0.140	0.654 \pm 0.020	86.418 \pm 6.841	0.995 \pm 0.001
	F-MIWAE	0.623 \pm 0.001	0.667 \pm 0.001	0.704 \pm 0.001	0.928 \pm 0.002	1.482 \pm 0.138	0.622 \pm 0.019	—	—
	F-NMIWAE	0.624 \pm 0.003	0.671 \pm 0.001	0.712 \pm 0.003	0.937 \pm 0.002	1.449 \pm 0.129	0.665 \pm 0.017	—	—
F-SAINT	F-GAIN	0.634 \pm 0.003	0.685 \pm 0.004	0.728 \pm 0.002	0.942 \pm 0.000	1.379 \pm 0.121	0.690 \pm 0.019	78.175 \pm 3.184	0.995 \pm 0.001
	F-Mean	0.637 \pm 0.003	0.688 \pm 0.002	0.723 \pm 0.002	0.941 \pm 0.001	1.323 \pm 0.001	0.721 \pm 0.001	72.491 \pm 5.497	0.996 \pm 0.001
	F-MIWAE	0.625 \pm 0.002	0.672 \pm 0.003	0.709 \pm 0.003	0.933 \pm 0.004	1.364 \pm 0.072	0.679 \pm 0.005	—	—
	F-NMIWAE	0.627 \pm 0.005	0.674 \pm 0.002	0.713 \pm 0.005	0.938 \pm 0.003	1.347 \pm 0.031	0.687 \pm 0.001	—	—
DARN	F-GAIN	0.644 \pm 0.006	0.702 \pm 0.004	0.734 \pm 0.002	0.947 \pm 0.003	1.291 \pm 0.104	0.748 \pm 0.011	67.198 \pm 1.487	0.997 \pm 0.001
	Central-DARN	0.643 \pm 0.001	0.700 \pm 0.003	0.727 \pm 0.002	0.943 \pm 0.002	1.335 \pm 0.033	0.693 \pm 0.007	74.164 \pm 7.928	0.997 \pm 0.001
	Local-DARN	0.639 \pm 0.001	0.696 \pm 0.001	0.721 \pm 0.002	0.940 \pm 0.003	1.362 \pm 0.013	0.677 \pm 0.004	71.948 \pm 6.156	0.997 \pm 0.001
	DARN	0.662 \pm 0.001	0.721 \pm 0.001	0.770 \pm 0.002	0.967 \pm 0.001	1.096 \pm 0.012	0.792 \pm 0.003	40.147 \pm 2.009	0.999 \pm 0.000

Table 2: The prediction performance under CCS.

Datasets		Higgs		Coverttype		Temp		Gas	
Models		Accuracy \uparrow	AUC \uparrow	Accuracy \uparrow	AUC \uparrow	RMSE \downarrow	R2 \uparrow	RMSE \downarrow	R2 \uparrow
F-XGBoost		0.651 \pm 0.005	0.710 \pm 0.005	0.703 \pm 0.002	0.929 \pm 0.003	1.257 \pm 0.019	0.738 \pm 0.008	99.681 \pm 4.651	0.994 \pm 0.001
F-GBDT		0.640 \pm 0.004	0.692 \pm 0.002	0.695 \pm 0.003	0.917 \pm 0.005	1.313 \pm 0.114	0.695 \pm 0.013	111.519 \pm 5.941	0.993 \pm 0.001
F-RF		0.621 \pm 0.001	0.662 \pm 0.002	0.668 \pm 0.006	0.891 \pm 0.003	—	—	—	—
F-MLP		0.631 \pm 0.005	0.682 \pm 0.004	0.689 \pm 0.003	0.902 \pm 0.004	1.515 \pm 0.169	0.612 \pm 0.029	127.581 \pm 3.654	0.992 \pm 0.001
F-TabNet		0.632 \pm 0.004	0.686 \pm 0.003	0.721 \pm 0.004	0.939 \pm 0.003	1.396 \pm 0.031	0.681 \pm 0.001	81.651 \pm 6.198	0.995 \pm 0.001
F-SAINT		0.643 \pm 0.003	0.701 \pm 0.002	0.732 \pm 0.001	0.944 \pm 0.002	1.315 \pm 0.016	0.714 \pm 0.020	71.948 \pm 4.738	0.997 \pm 0.001
Central-DARN		0.641 \pm 0.001	0.694 \pm 0.003	0.727 \pm 0.003	0.941 \pm 0.004	1.323 \pm 0.052	0.670 \pm 0.009	68.417 \pm 6.185	0.997 \pm 0.001
Local-DARN		0.637 \pm 0.002	0.689 \pm 0.002	0.720 \pm 0.003	0.939 \pm 0.002	1.367 \pm 0.036	0.641 \pm 0.004	76.779 \pm 3.617	0.997 \pm 0.001
DARN		0.658 \pm 0.002	0.717 \pm 0.001	0.767 \pm 0.002	0.964 \pm 0.002	1.183 \pm 0.016	0.771 \pm 0.003	45.164 \pm 3.698	0.999 \pm 0.001

Table 3: The prediction performance under PCS.

Datasets		Higgs		Coverttype		Temp		Gas	
Models		Accuracy \uparrow	AUC \uparrow	Accuracy \uparrow	AUC \uparrow	RMSE \downarrow	R2 \uparrow	RMSE \downarrow	R2 \uparrow
F-XGBoost		0.649 \pm 0.003	0.708 \pm 0.002	0.701 \pm 0.001	0.926 \pm 0.002	1.279 \pm 0.074	0.736 \pm 0.005	107.982 \pm 0.669	0.994 \pm 0.001
F-GBDT		0.636 \pm 0.004	0.688 \pm 0.003	0.695 \pm 0.002	0.915 \pm 0.002	1.333 \pm 0.021	0.687 \pm 0.003	114.648 \pm 1.233	0.994 \pm 0.001
F-RF		0.622 \pm 0.001	0.661 \pm 0.001	0.664 \pm 0.003	0.886 \pm 0.001	—	—	—	—
F-MLP		0.626 \pm 0.001	0.676 \pm 0.001	0.684 \pm 0.001	0.894 \pm 0.001	1.545 \pm 0.054	0.607 \pm 0.004	131.495 \pm 2.541	0.992 \pm 0.001
F-TabNet		0.634 \pm 0.005	0.684 \pm 0.004	0.724 \pm 0.003	0.940 \pm 0.001	1.413 \pm 0.140	0.676 \pm 0.020	86.176 \pm 2.481	0.995 \pm 0.001
F-SAINT		0.636 \pm 0.001	0.693 \pm 0.002	0.728 \pm 0.001	0.942 \pm 0.002	1.328 \pm 0.015	0.706 \pm 0.004	73.486 \pm 4.561	0.997 \pm 0.001
Central-DARN		0.640 \pm 0.004	0.693 \pm 0.004	0.725 \pm 0.003	0.941 \pm 0.002	1.328 \pm 0.008	0.667 \pm 0.001	70.165 \pm 7.169	0.997 \pm 0.001
Local-DARN		0.637 \pm 0.002	0.690 \pm 0.003	0.714 \pm 0.002	0.936 \pm 0.002	1.371 \pm 0.039	0.636 \pm 0.005	79.146 \pm 5.532	0.997 \pm 0.001
DARN		0.653 \pm 0.001	0.713 \pm 0.002	0.754 \pm 0.001	0.957 \pm 0.002	1.215 \pm 0.027	0.765 \pm 0.001	52.194 \pm 5.024	0.999 \pm 0.000

Table 4: The prediction performance under SSCS.

Datasets		Higgs		Coverttype		Temp		Gas	
Models		Accuracy \uparrow	AUC \uparrow	Accuracy \uparrow	AUC \uparrow	RMSE \downarrow	R2 \uparrow	RMSE \downarrow	R2 \uparrow
F-XGBoost		0.646 \pm 0.002	0.706 \pm 0.002	0.689 \pm 0.006	0.913 \pm 0.005	1.284 \pm 0.100	0.732 \pm 0.010	109.781 \pm 5.517	0.994 \pm 0.001
F-GBDT		0.634 \pm 0.007	0.687 \pm 0.010	0.684 \pm 0.005	0.913 \pm 0.003	1.376 \pm 0.147	0.675 \pm 0.022	117.982 \pm 1.981	0.993 \pm 0.001
F-RF		0.618 \pm 0.008	0.656 \pm 0.006	0.651 \pm 0.002	0.871 \pm 0.003	—	—	—	—
F-MLP		0.623 \pm 0.003	0.671 \pm 0.002	0.673 \pm 0.007	0.886 \pm 0.004	1.549 \pm 0.092	0.589 \pm 0.008	135.714 \pm 4.897	0.992 \pm 0.001
F-TabNet		0.628 \pm 0.003	0.675 \pm 0.001	0.710 \pm 0.001	0.931 \pm 0.002	1.428 \pm 0.084	0.676 \pm 0.007	90.641 \pm 3.983	0.994 \pm 0.001
F-SAINT		0.638 \pm 0.001	0.695 \pm 0.001	0.722 \pm 0.001	0.936 \pm 0.002	1.346 \pm 0.080	0.701 \pm 0.06	76.415 \pm 6.614	0.997 \pm 0.001
Central-DARN		0.643 \pm 0.001	0.697 \pm 0.002	0.726 \pm 0.004	0.941 \pm 0.005	1.291 \pm 0.041	0.721 \pm 0.011	64.517 \pm 2.148	0.998 \pm 0.000
Local-DARN		0.638 \pm 0.002	0.693 \pm 0.003	0.717 \pm 0.001	0.938 \pm 0.003	1.351 \pm 0.038	0.689 \pm 0.005	68.492 \pm 4.738	0.997 \pm 0.001
DARN		0.648 \pm 0.001	0.707 \pm 0.001	0.731 \pm 0.001	0.947 \pm 0.001	1.259 \pm 0.046	0.754 \pm 0.014	55.134 \pm 6.517	0.999 \pm 0.001

Table 5: The prediction performance under CRS.

Datasets	Higgs		Coverttype		Temp		Gas	
Models	Accuracy \uparrow	AUC \uparrow	Accuracy \uparrow	AUC \uparrow	RMSE \downarrow	R2 \uparrow	RMSE \downarrow	R2 \uparrow
F-XGBoost	0.642 \pm 0.001	0.698 \pm 0.002	0.705 \pm 0.003	0.929 \pm 0.002	1.289 \pm 0.037	0.729 \pm 0.004	105.564 \pm 3.477	0.994 \pm 0.001
F-GBDT	0.632 \pm 0.006	0.687 \pm 0.02	0.696 \pm 0.006	0.918 \pm 0.004	1.344 \pm 0.016	0.701 \pm 0.002	103.189 \pm 3.655	0.993 \pm 0.001
F-RF	0.618 \pm 0.003	0.655 \pm 0.004	0.676 \pm 0.001	0.901 \pm 0.002	—	—	—	—
F-MLP	0.626 \pm 0.004	0.673 \pm 0.005	0.688 \pm 0.005	0.899 \pm 0.002	1.514 \pm 0.062	0.621 \pm 0.011	114.487 \pm 7.246	0.993 \pm 0.001
F-TabNet	0.631 \pm 0.005	0.675 \pm 0.003	0.719 \pm 0.001	0.936 \pm 0.003	1.352 \pm 0.028	0.693 \pm 0.003	82.791 \pm 4.489	0.994 \pm 0.001
F-SAINT	0.641 \pm 0.001	0.698 \pm 0.001	0.737 \pm 0.002	0.948 \pm 0.003	1.273 \pm 0.012	0.741 \pm 0.003	69.486 \pm 3.332	0.997 \pm 0.001
Central-DARN	0.638 \pm 0.003	0.692 \pm 0.002	0.733 \pm 0.004	0.946 \pm 0.003	1.271 \pm 0.025	0.739 \pm 0.008	71.912 \pm 6.166	0.998 \pm 0.001
Local-DARN	0.636 \pm 0.004	0.688 \pm 0.003	0.723 \pm 0.002	0.939 \pm 0.004	1.385 \pm 0.072	0.671 \pm 0.017	73.984 \pm 8.728	0.997 \pm 0.001
DARN	0.660 \pm 0.002	0.720 \pm 0.001	0.773 \pm 0.001	0.971 \pm 0.001	1.138 \pm 0.009	0.783 \pm 0.001	38.624 \pm 1.137	0.999 \pm 0.000

Table 6: The prediction performance under all random scenarios.

Datasets	Higgs		Coverttype		Temp		Gas	
Models	Accuracy \uparrow	AUC \uparrow	Accuracy \uparrow	AUC \uparrow	RMSE \downarrow	R2 \uparrow	RMSE \downarrow	R2 \uparrow
F-XGBoost	0.628 \pm 0.003	0.683 \pm 0.002	0.691 \pm 0.004	0.914 \pm 0.003	1.312 \pm 0.040	0.715 \pm 0.005	108.245 \pm 3.500	0.993 \pm 0.002
F-GBDT	0.619 \pm 0.005	0.672 \pm 0.003	0.682 \pm 0.005	0.903 \pm 0.004	1.368 \pm 0.020	0.689 \pm 0.003	115.892 \pm 3.700	0.993 \pm 0.002
F-RF	0.605 \pm 0.002	0.641 \pm 0.003	0.662 \pm 0.003	0.886 \pm 0.002	—	—	—	—
F-MLP	0.613 \pm 0.003	0.659 \pm 0.004	0.673 \pm 0.004	0.892 \pm 0.003	1.587 \pm 0.065	0.605 \pm 0.012	132.782 \pm 7.300	0.991 \pm 0.002
F-TabNet	0.621 \pm 0.006	0.661 \pm 0.005	0.705 \pm 0.003	0.921 \pm 0.003	1.452 \pm 0.150	0.669 \pm 0.022	87.345 \pm 4.600	0.994 \pm 0.002
F-SAINT	0.625 \pm 0.002	0.683 \pm 0.003	0.713 \pm 0.002	0.931 \pm 0.003	1.362 \pm 0.025	0.699 \pm 0.005	74.123 \pm 4.800	0.996 \pm 0.002
Central-DARN	0.623 \pm 0.004	0.681 \pm 0.004	0.709 \pm 0.004	0.929 \pm 0.003	1.360 \pm 0.015	0.690 \pm 0.003	72.345 \pm 7.300	0.996 \pm 0.002
Local-DARN	0.620 \pm 0.003	0.678 \pm 0.004	0.698 \pm 0.003	0.924 \pm 0.004	1.405 \pm 0.075	0.660 \pm 0.010	81.234 \pm 8.800	0.996 \pm 0.002
DARN	0.645 \pm 0.002	0.705 \pm 0.003	0.758 \pm 0.002	0.956 \pm 0.003	1.225 \pm 0.035	0.758 \pm 0.002	51.456 \pm 5.200	0.998 \pm 0.001

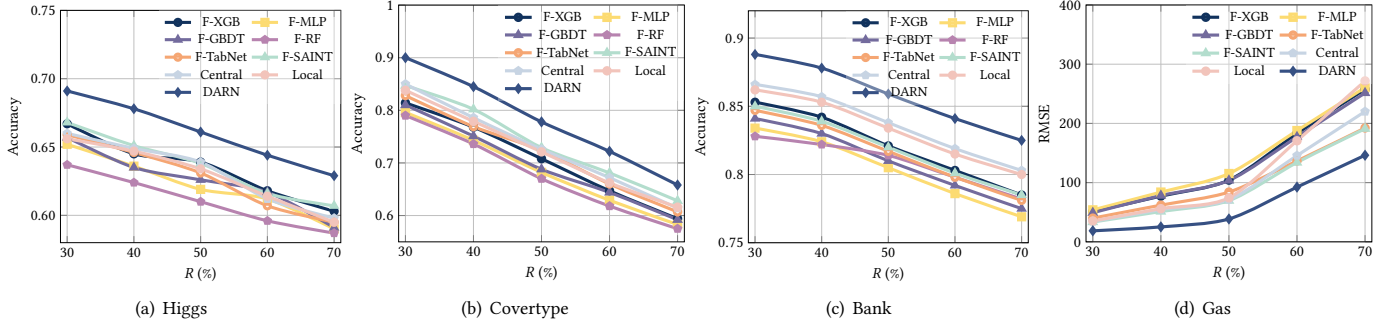


Figure 1: The prediction performance of tabular data prediction algorithms vs. missing rate R .

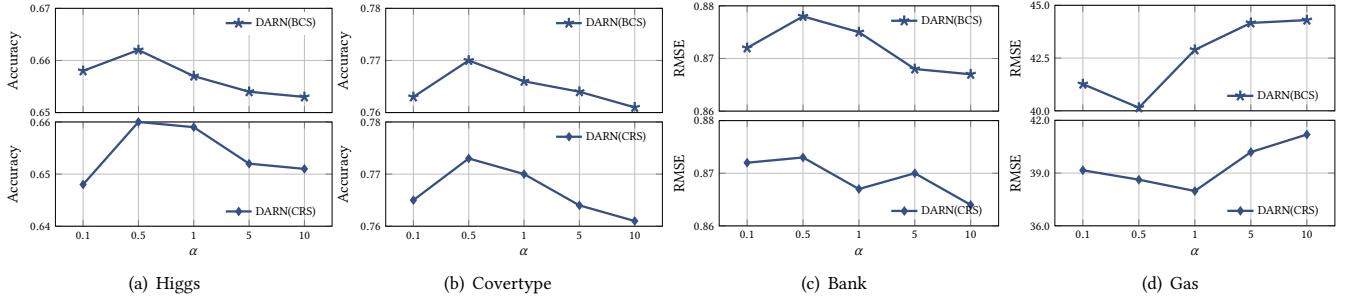


Figure 2: The prediction performance of DARN vs. weight hyperparameter α .