

## Experimental Results

Dataset	Observations	Features
cpu_act	8,192	22
pol	15,000	27
elevators	16,599	17
isolet	7,797	614
wine_quality	6,497	12
Ailerons	13,750	34
houses	20,640	9
house_16H	22,784	17
diamonds	53,940	7
Brazilian_houses	10,692	9
Bike_Sharing_Demand	17,379	7
nyc-taxi-green-dec-2016	581,835	10
house_sales	21,613	16
sulfur	10,081	7
medical_charges	163,065	4
MiamiHousing2016	13,932	14
superconduct	21,263	80
california	20,640	9
fifa	18,063	6
year	515,345	91

Table 1: Overview of regression datasets including the number of observations and features.

Table 2: Regression Results (MSE)

Dataset	GF	RF	XGB	LGBM	Diff.
cpu act	<b>5.25</b> (0.11)	10.40 (0.14)	5.37 (0.14)	5.82 (0.26)	2.23%
pol	<b>20.93</b> (0.56)	132.61 (2.05)	42.90 (1.08)	22.42 (0.54)	6.65%
elevators ( $10^{-6}$ )	<b>4.90</b> (0.05)	14.86 (0.16)	5.45 (0.07)	4.99 (0.06)	1.8%
isolet	—	—	—	—	—
wine quality	—	—	—	—	—
Wine Quality	—	—	—	—	—
Wine Quality	—	—	—	—	—

Table 3: Regression results for major tree-based models. The best result for each dataset is highlighted in bold. The final column shows the percentage difference between GF and the best non-GF model. Std of the mean result is shown in parenthesis.

Table 4: Classification Results (Log-Loss)

Dataset	Obs.	Dim.	GF	RF	XGB	LGBM	CatBoost
10M Higgs	11M	28	—	—	—	—	—
Cardiovascular Disease	—	11	—	—	—	—	—
Heart Disease	—	—	—	—	—	—	—
Airline Satisfaction	—	—	—	—	—	—	—
Health Insurance Interest	—	—	—	—	—	—	—
L&T Vehicle Loan	—	—	—	—	—	—	—
Wine Quality	—	—	—	—	—	—	—
Wine Quality	—	—	—	—	—	—	—
Wine Quality	—	—	—	—	—	—	—
Wine Quality	—	—	—	—	—	—	—
Wine Quality	—	—	—	—	—	—	—
Wine Quality	—	—	—	—	—	—	—