Table 1: Performance on more metrics compared with FeatLLM.

		acc		precision		recall		f1_score	
dataset	shot	FeatLLM	${\bf ProtoLLM}$	FeatLLM	ProtoLLM	FeatLLM	${\bf ProtoLLM}$	FeatLLM	ProtoLLM
Adult	4	44.82±24.99	$56.24{\pm}5.04$	63.36±16.94	$66.84{\pm}1.09$	57.89±9.49	$70.35{\pm}2.85$	38.77±21.95	$55.77{\pm}4.65$
	8	48.24 ± 20.09	$58.75 {\pm} 4.84$	53.38 ± 24.08	$67.33{\pm}1.08$	64.52 ± 11.75	$71.75{\pm}2.66$	45.69 ± 21.04	$58.05{\pm}4.35$
	16	$64.95{\pm}18.48$	$62.86{\pm}4.58$	$71.42{\pm}6.46$	$68.25{\pm}1.13$	68.34 ± 9.21	$73.97 {\pm} 2.42$	58.84 ± 15.99	$61.70{\pm}4.02$
Bank	4	$32.42{\pm}13.68$	$73.21 {\pm} 6.48$	55.40 ± 1.12	$61.31{\pm}1.86$	58.32 ± 4.73	$72.55{\pm}3.60$	30.66 ± 10.73	$60.72{\pm}4.08$
	8	38.50 ± 12.31	$75.26 {\pm} 3.93$	55.72 ± 0.76	$61.72 {\pm} 1.69$	60.57 ± 4.47	$73.04{\pm}3.09$	35.69 ± 10.40	$62.17 {\pm} 2.82$
	16	$45.21{\pm}10.61$	$76.67 {\pm} 3.83$	56.35 ± 0.96	$62.92 {\pm} 1.38$	63.57 ± 3.65	$75.07 {\pm} 1.75$	41.25 ± 7.75	$63.92 {\pm} 2.65$
Blood	4	62.94 ± 7.12	63.11 ± 9.33	60.73±3.94	$62.86{\pm}5.37$	63.31±4.29	$65.55{\pm}6.81$	58.55±5.35	58.74 ± 7.57
	8	$66.36{\pm}4.89$	$64.98{\pm}6.98$	60.85 ± 3.39	$63.60{\pm}5.10$	62.96 ± 5.46	$67.67{\pm}6.42$	59.31 ± 3.92	$61.25{\pm}6.58$
	16	64.06 ± 8.43	$65.29{\pm}4.88$	60.60 ± 3.78	$63.20{\pm}3.52$	62.21 ± 4.73	$67.43{\pm}4.43$	57.85 ± 6.10	$61.65{\pm}4.38$
Credit-g	4	64.97 ± 4.43	$69.97{\pm}2.97$	51.13±7.60	$64.35{\pm}10.74$	51.01±2.55	$51.77{\pm}1.45$	47.33±4.63	46.23±2.84
	8	$64.13{\pm}4.97$	$69.90{\pm}3.32$	52.76 ± 5.10	$63.59 {\pm} 9.91$	51.59 ± 3.15	$52.72{\pm}1.93$	50.39 ± 3.59	48.56 ± 3.79
	16	$63.43{\pm}4.68$	$69.67{\pm}3.04$	55.71 ± 6.59	$61.11 {\pm} 3.87$	54.39 ± 5.12	$55.48{\pm}2.32$	53.03 ± 5.84	$54.23{\pm}3.45$

Table 2: Performance on Weighted ProtoLLM and Weighted ProtoLLM with half feature droped.

dataset	shot	Weighted-protoLLM	Weighted-protoLLM(del features)
	4	85.34	85.32
adult	8	85.5	85.45
0.04.01.0	16	85.72	85.63
	4	81.01	80.4
bank	8	81.01	80.35
	16	81.29	80.63
	4	71.34	69.59
blood	8	71.46	69.19
	16	72.85	70.26
	4	80.73	80.42
diabetes	8	80.79	80.47
	16	80.72	80.29
	4	95.95	96.43
NHANES	8	96.57	96.92
	16	98.37	98.47

Table 3: Performance of ProtoLLM with poor feature descriptions or without feature descriptions..

dataset	shot	full-desc	poor-desc	no-desc
	4	86.01 ± 0.78	$83.36 {\pm} 0.67$	$83.63 {\pm} 0.55$
Adult	8	86.12 ± 0.92	83.68 ± 0.67	83.90 ± 0.59
	16	86.28 ± 0.77	83.91 ± 0.82	84.08 ± 0.81
	4	80.85 ± 2.58	71.99 ± 1.93	61.15 ± 4.91
Bank	8	81.41 ± 2.58	72.98 ± 1.90	$62.86{\pm}4.98$
	16	83.26 ± 1.40	75.39 ± 1.65	65.53 ± 4.67
	4	75.98 ± 4.99	72.78 ± 7.12	62.44 ± 9.80
Blood	8	76.35 ± 4.61	74.69 ± 6.58	64.18 ± 8.88
	16	75.46 ± 4.12	$74.88{\pm}4.75$	$66.54{\pm}5.84$
	4	62.25±2.86	51.32 ± 4.75	51.39 ± 4.85
Credit-g	8	63.26 ± 2.87	52.34 ± 4.54	52.74 ± 4.27
	16	64.52 ± 3.28	54.59 ± 4.66	55.48 ± 4.73