Table 1: Rescaled Squentropy Performance on ViT and CIFAR-10 compared to Cross-Entropy

$$l_{\text{sqen}}(\vec{x_i}, y_i) = -\log p_{i,y_i}(\vec{x_i}) + \frac{\alpha}{C-1} \sum_{j=1,\, j\neq y_i}^{C} f_j(\vec{x_i})^2$$

Domain	Model	Task	Squentropy					Cross-Entropy		
			lr	alpha	inc class	perf	ECE	lr	perf	ECE
Vision	ViT	CIFAR-10	2×10^{-4}	1.0	0	82.78	3.90	2×10^{-4}	84.34	3.56
						82.95	3.62		84.02	3.19
						82.41	3.58		84.35	3.18
						82.82	3.50		84.42	3.17
						83.02	3.45		84.02	3.14
					avg	82.80	3.61	avg	84.23	3.25
			2×10^{-4}	1.0	-1	82.34	4.01			
						82.32	3.84			
						81.81	3.60			
						82.34	3.49			
						82.23	3.26			
					avg	82.21	3.64			
			2×10^{-4}	0.1	0	83.97	2.85			
						84.42	2.67			
						84.07	2.65			
						84.06	2.49			
						84.15	2.37			
					avg	84.13	2.61			
			2×10^{-4}	0.1	-1	84.42	2.42			
						84.73	2.42			
						84.55	2.40			
						84.19	2.32			
						84.46	2.27			
					avg	84.47	2.37			
			2×10^{-4}	0.25	0	83.54	2.44			
						84.04	2.43			
						83.92	1.90			
					avg	83.83	2.26			
			2×10^{-4}	0.5	0	83.11	2.41			
						83.11	2.04			
						83.23	1.93			
					avg	83.15	2.13			