

Temperature Scaling

Softmax with Temperature

$$p_i = \exp(z_i/T) / \sum \exp(z_j/T)$$

T = Temperature 파라미터



T = 1

일반 Softmax

[0.05, 0.10, 0.80, 0.05]

Sharp 분포



T = 3~5

종류에 최적

[0.15, 0.20, 0.45, 0.20]

Soft 분포



T → ∞

극한 케이스

[0.25, 0.25, 0.25, 0.25]

Uniform 분포



$T=1$

$T=3\sim 5$

$T \rightarrow \infty$



Sharp

Optimal

Uniform

실전 팁: 의료 영상 분류에서 $T=3\sim 4$ 가 효과적
높은 $T \rightarrow$ 클래스 간 관계 정보 풍부 \rightarrow **Student** 학습 향상