

Bias correction factor $\frac{\sqrt{1-\beta_2^t}}{(1-\beta_1^t)} * \text{warmup_cosine}(t)$

