Vector Training Losses - Llama 3.1 8B [adaptive_linear]