Distribution of "attention_key" Weights in microsoft/DialoGPT-small Weight Distribution Std. Dev: 0.1369 Mean: -0.0000 Kurtosis: 17.9378 Count: 7,077,888 3.0 2.5 2.0 Density 1.5 1.0 0.5 0.0 _ __2 -3 -1

Weight Value