Distributions for "attention_value" in All Models (Sample: HuggingFaceTB/SmolLM-360M)

Distribution of Weights