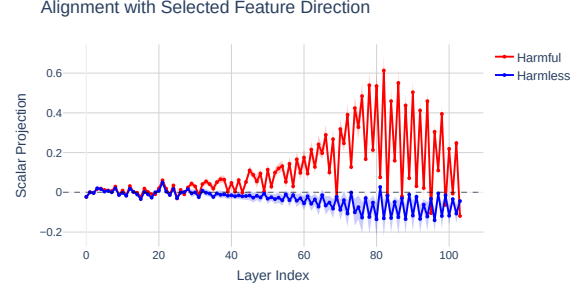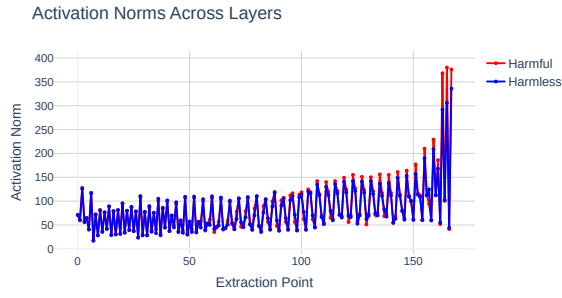(a) Activation norms across layers

(b) Projections on local candidate directions

Figure 5: **Layer-wise heterogeneity in gemma-2-2b-it.**

(a) Activation norms across layers

(b) Projections on local candidate directions

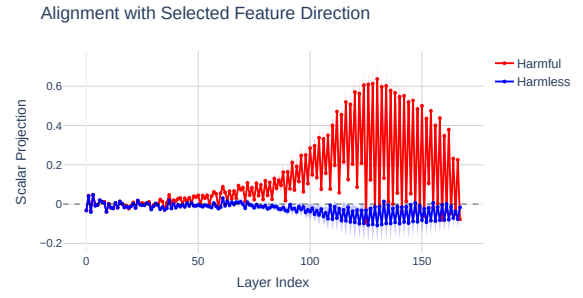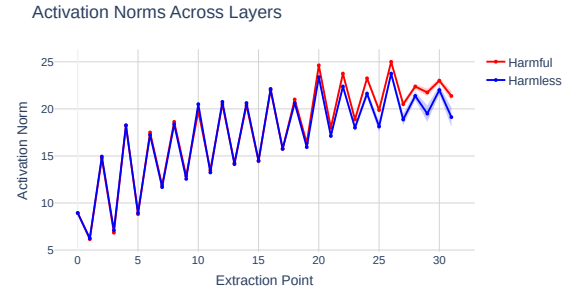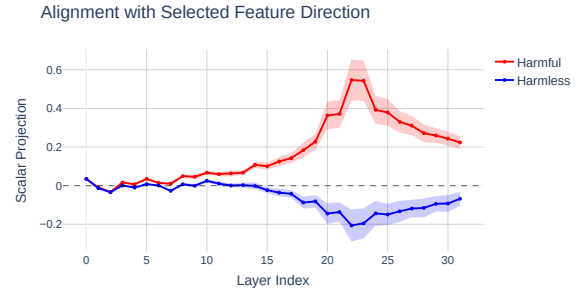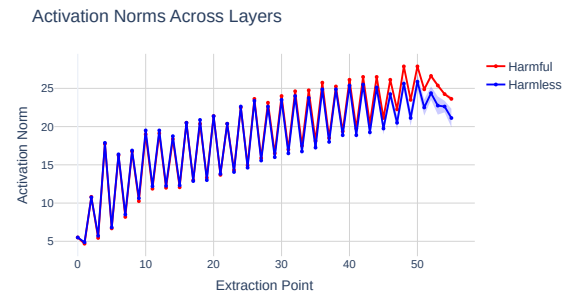Figure 6: **Layer-wise heterogeneity in gemma-2-9b-it.**
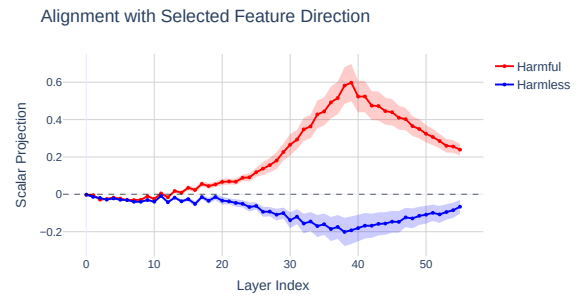

(a) Activation norms across layers

(b) Projections on local candidate directions

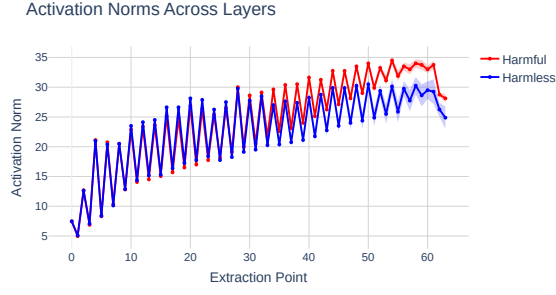Figure 7: **Layer-wise heterogeneity in Llama-3.2-1B-Instruct.**

(a) Activation norms across layers

(b) Projections on local candidate directions

Figure 8: **Layer-wise heterogeneity in Llama-3.2-3B-Instruct.**

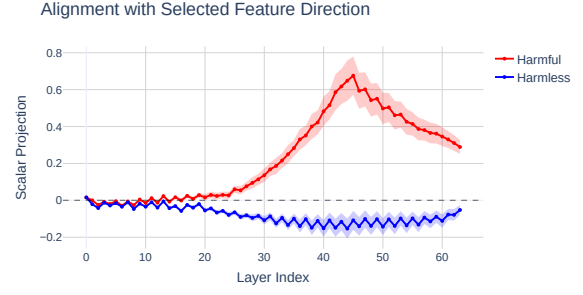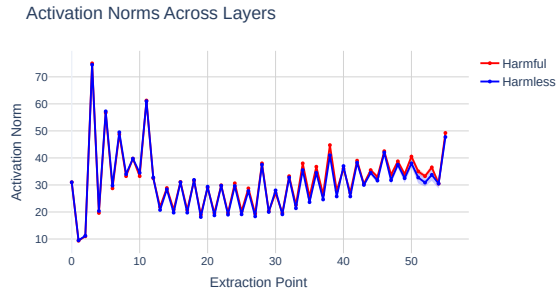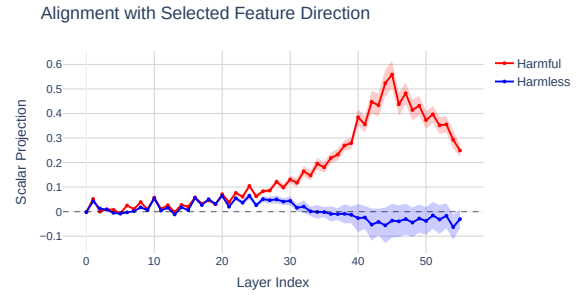(a) Activation norms across layers

(b) Projections on local candidate directions

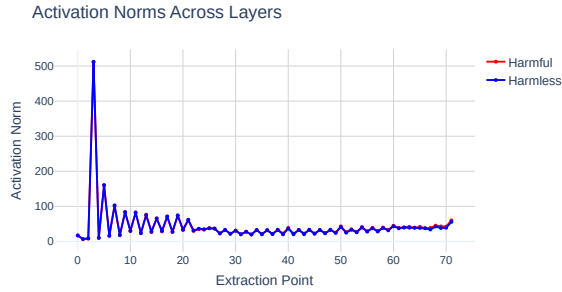Figure 9: **Layer-wise heterogeneity in Llama-3.1-8B-Instruct.**
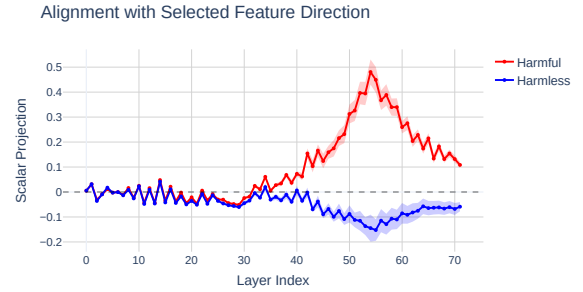

(a) Activation norms across layers

(b) Projections on local candidate directions

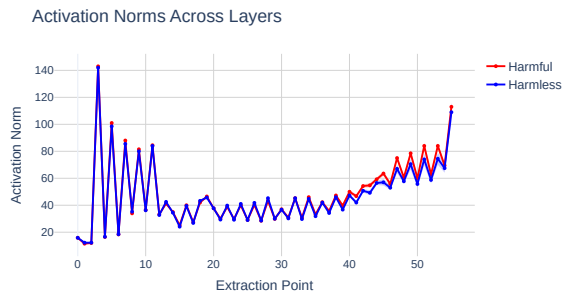Figure 10: **Layer-wise heterogeneity in Qwen2.5-1.5B-Instruct.**
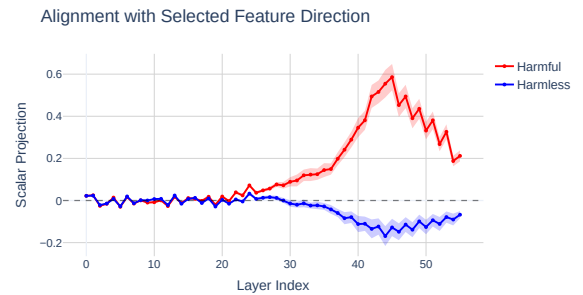

(a) Activation norms across layers

(b) Projections on local candidate directions

Figure 11: **Layer-wise heterogeneity in Qwen2.5-3B-Instruct.**

(a) Activation norms across layers

(b) Projections on local candidate directions

Figure 12: **Layer-wise heterogeneity in Qwen2.5-7B-Instruct.**