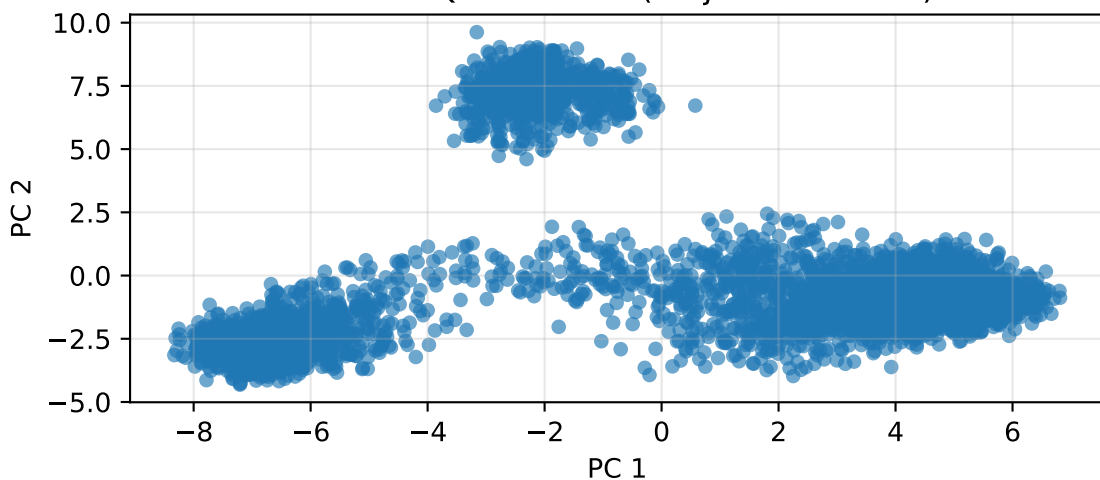
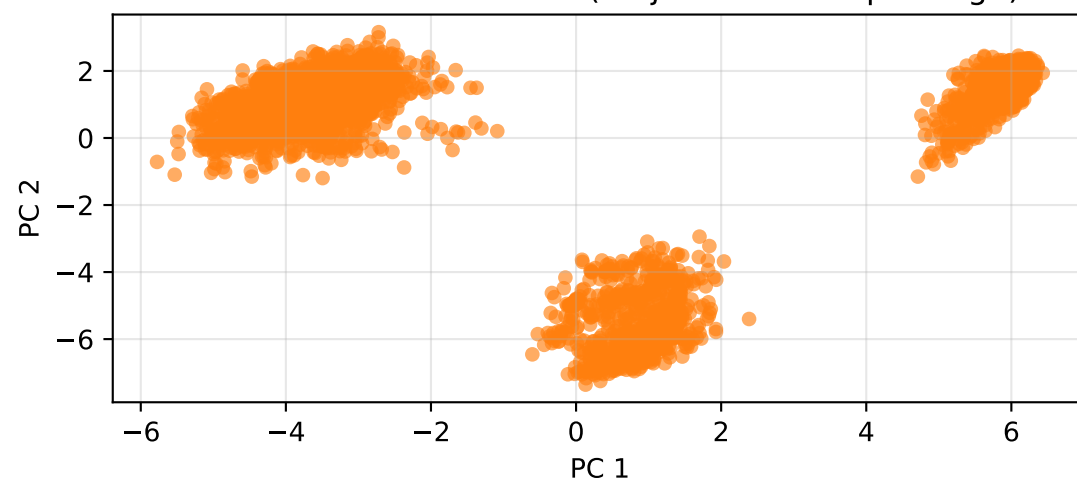


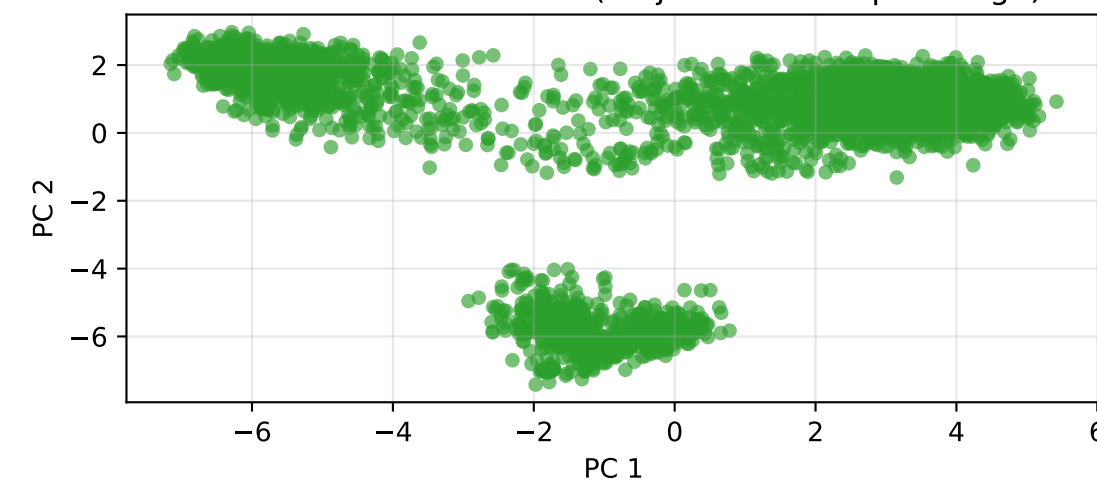
ATTN — Qwen2.5-7B (Projected Trainer)



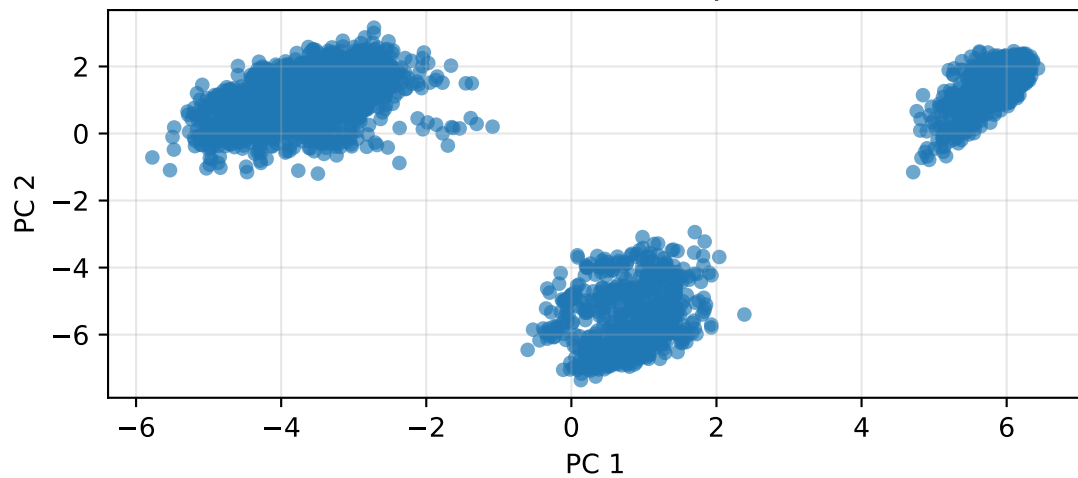
ATTN — Falcon3-7B-Base (Projected Tester pre-align)



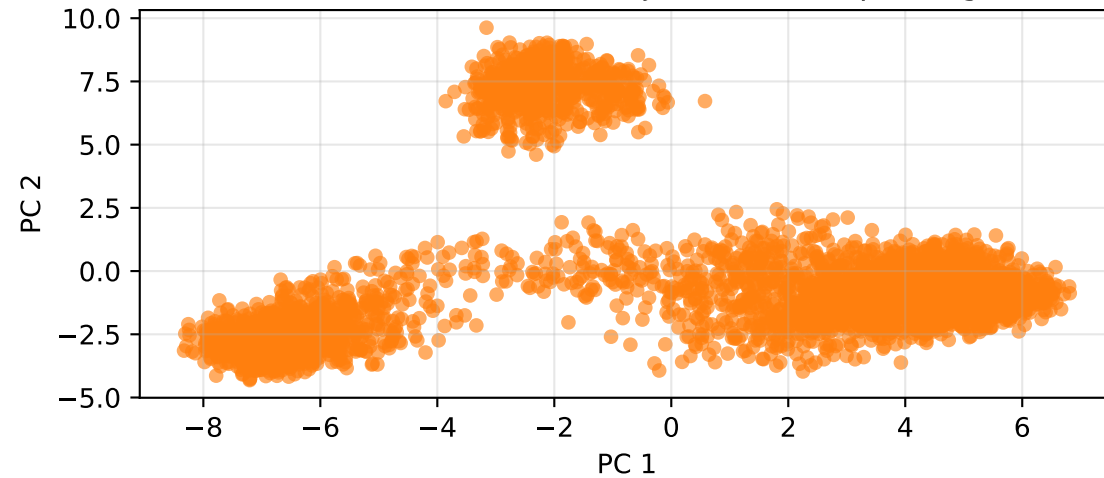
ATTN — Falcon3-7B-Base (Projected Tester post-align)



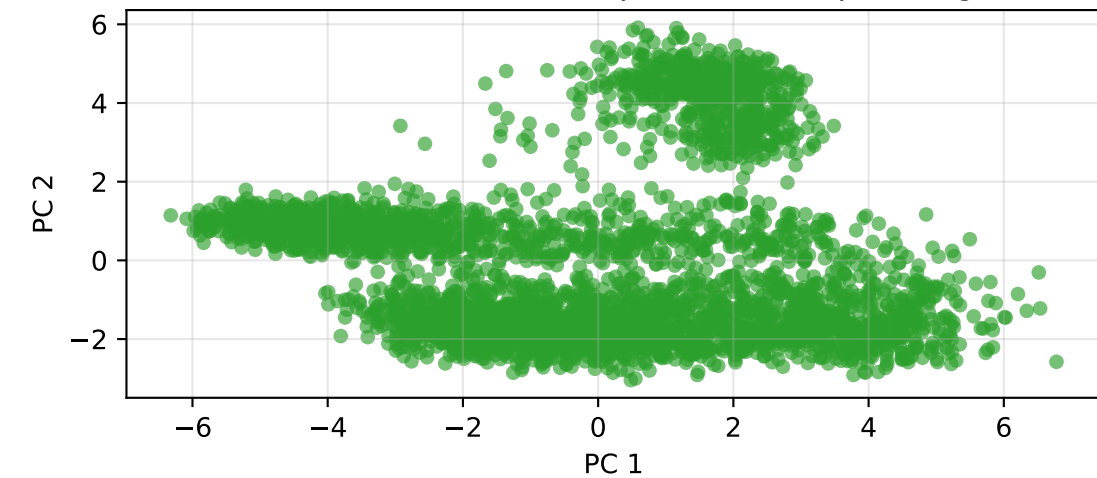
ATTN — Falcon3-7B-Base (Projected Trainer)



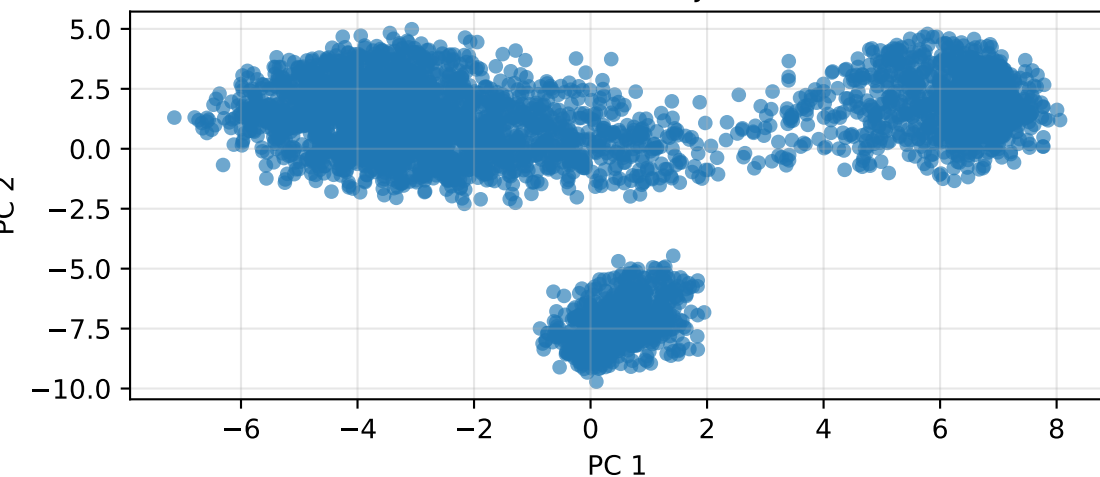
ATTN — Qwen2.5-7B (Projected Tester pre-align)



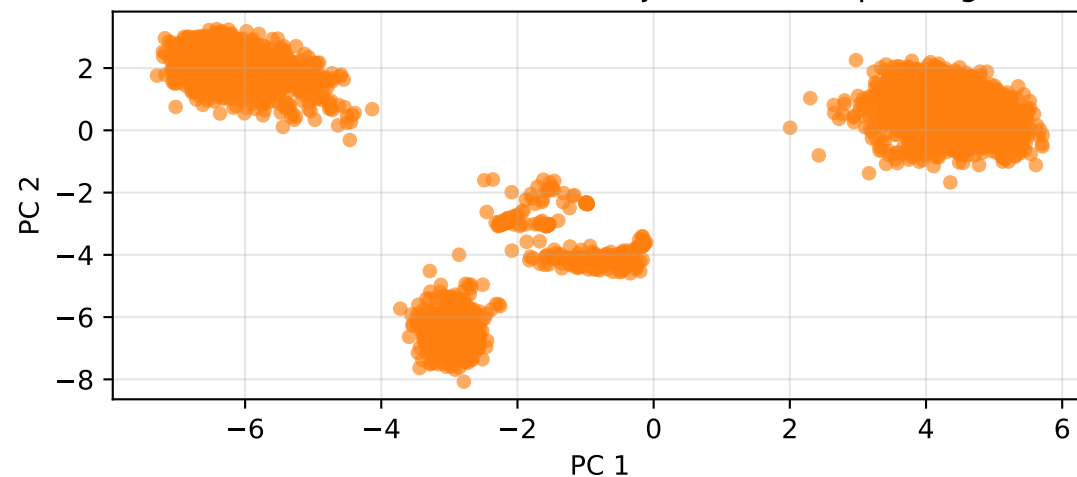
ATTN — Qwen2.5-7B (Projected Tester post-align)



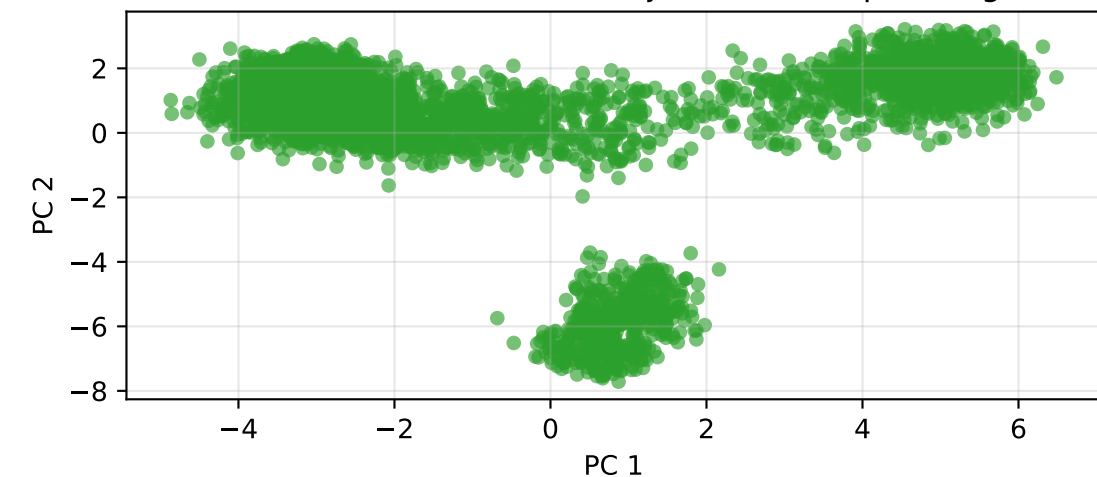
MLP — Qwen2.5-7B (Projected Trainer)



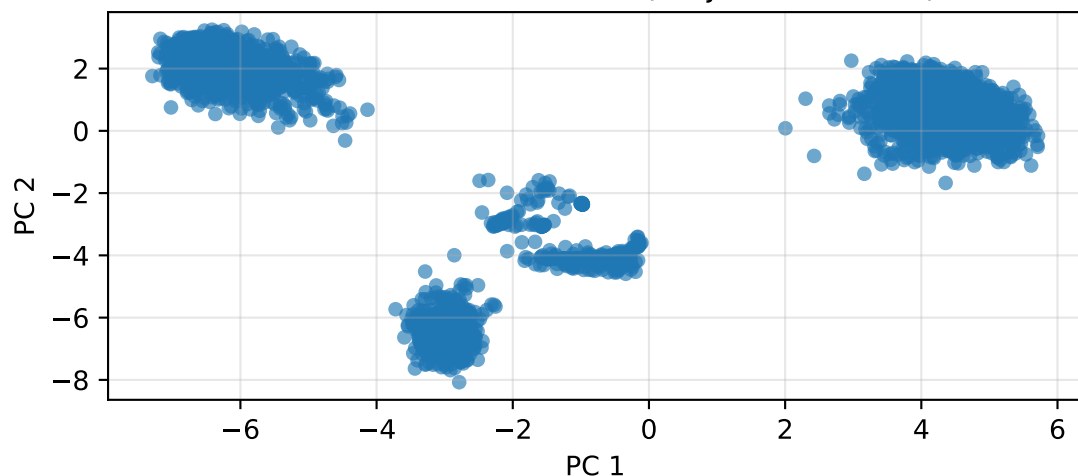
MLP — Falcon3-7B-Base (Projected Tester pre-align)



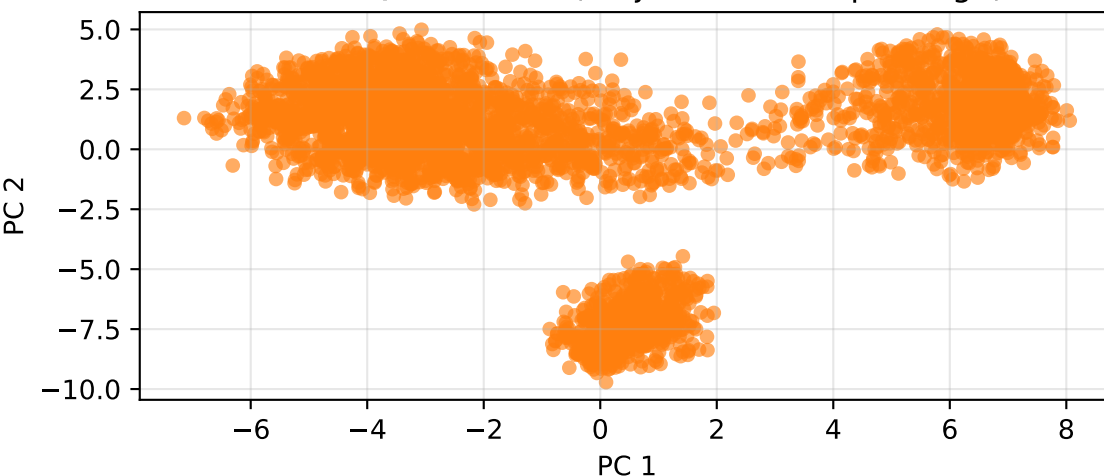
MLP — Falcon3-7B-Base (Projected Tester post-align)



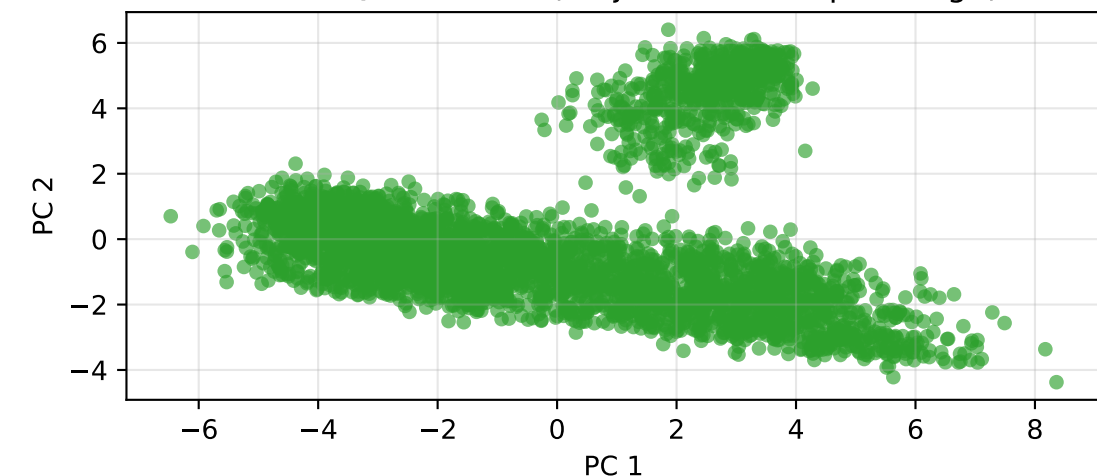
MLP — Falcon3-7B-Base (Projected Trainer)



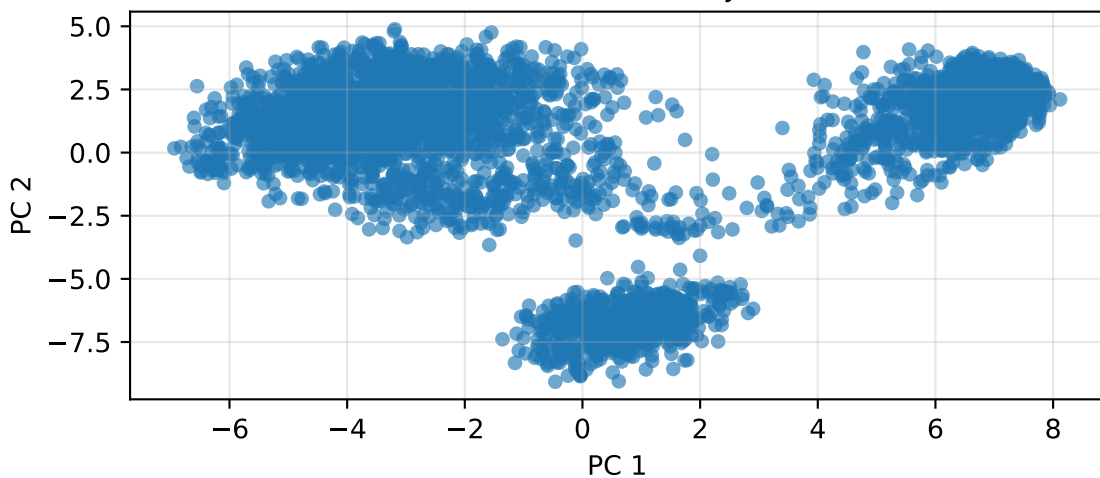
MLP — Qwen2.5-7B (Projected Tester pre-align)



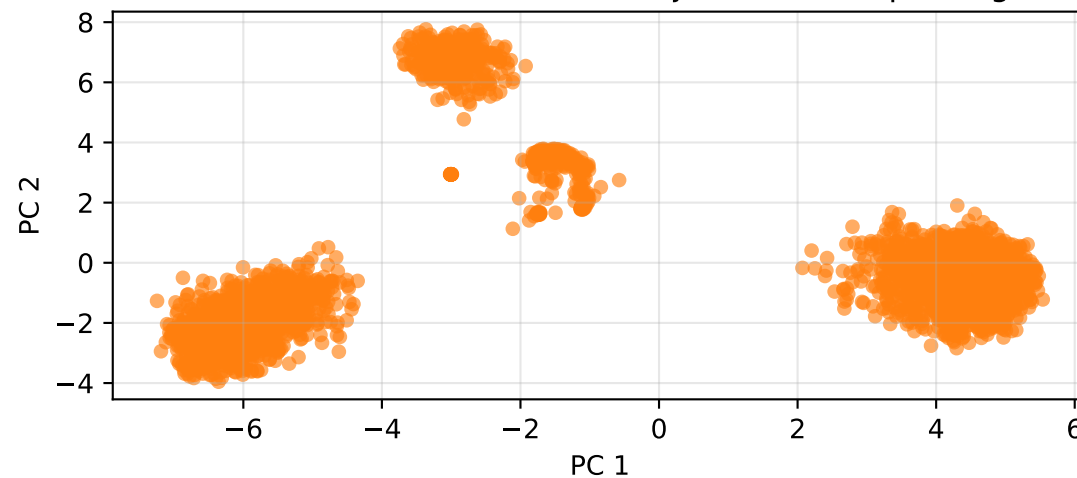
MLP — Qwen2.5-7B (Projected Tester post-align)



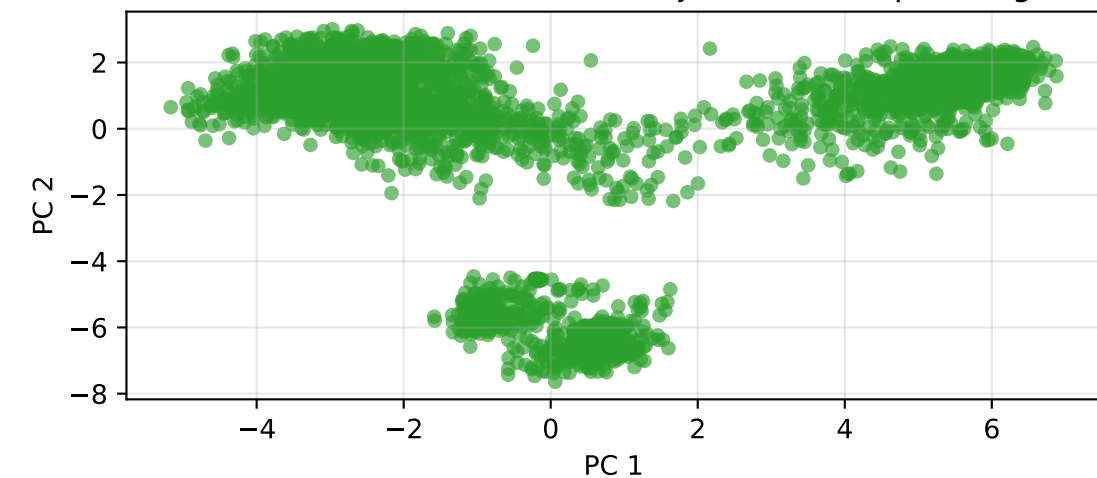
HIDDEN — Qwen2.5-7B (Projected Trainer)



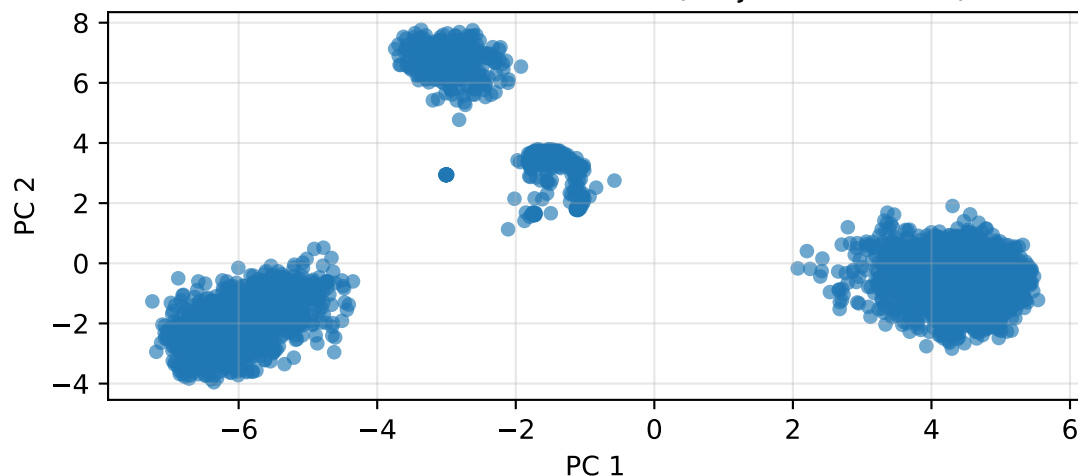
HIDDEN — Falcon3-7B-Base (Projected Tester pre-align)



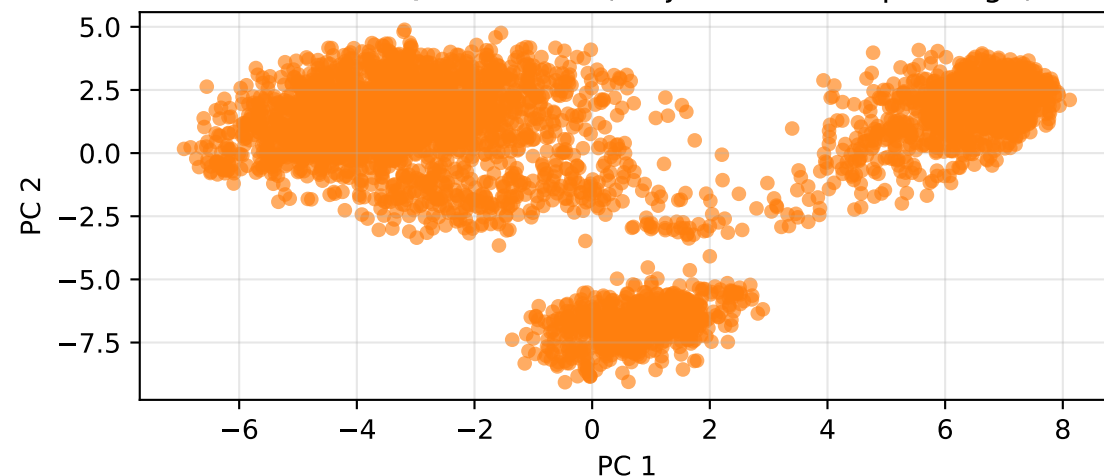
HIDDEN — Falcon3-7B-Base (Projected Tester post-align)



HIDDEN — Falcon3-7B-Base (Projected Trainer)



HIDDEN — Qwen2.5-7B (Projected Tester pre-align)



HIDDEN — Qwen2.5-7B (Projected Tester post-align)

