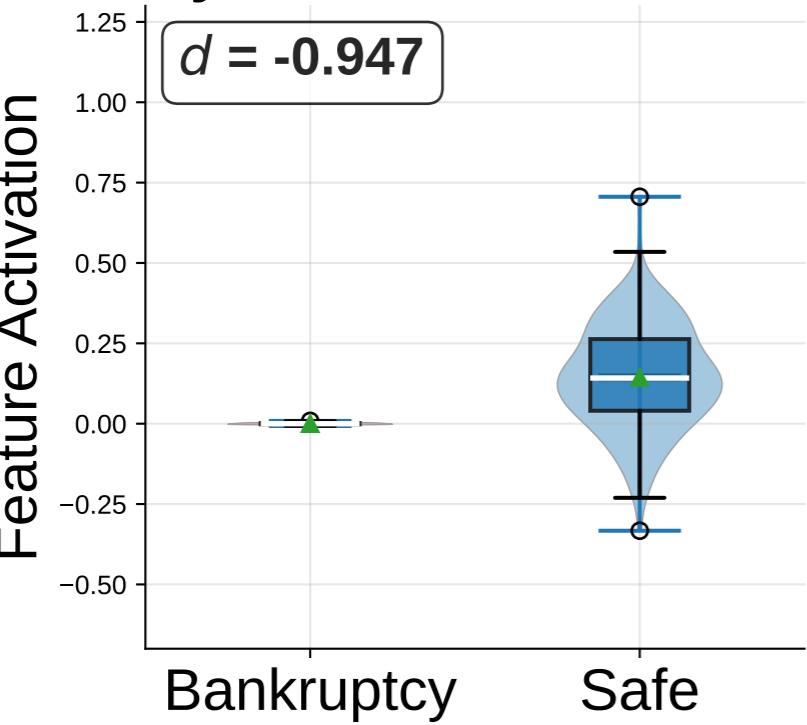
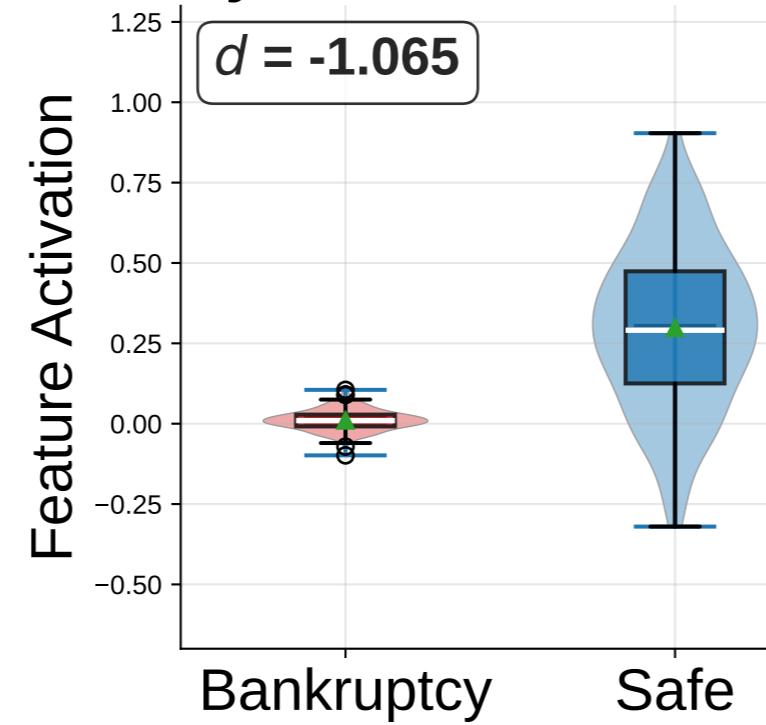


SAE Feature Activation Distributions: Bankruptcy vs Safe Groups

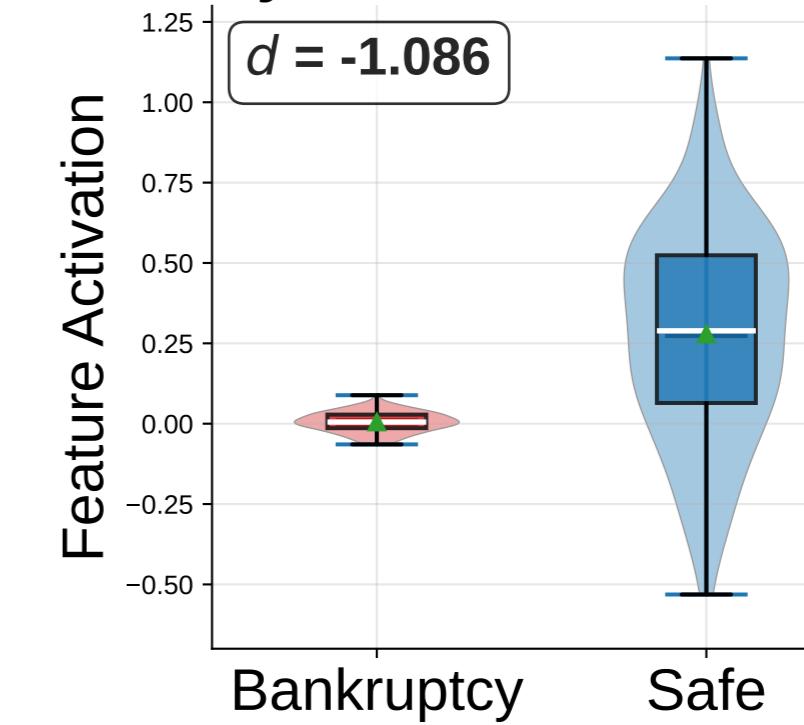
Layer 25 - Feature 17793



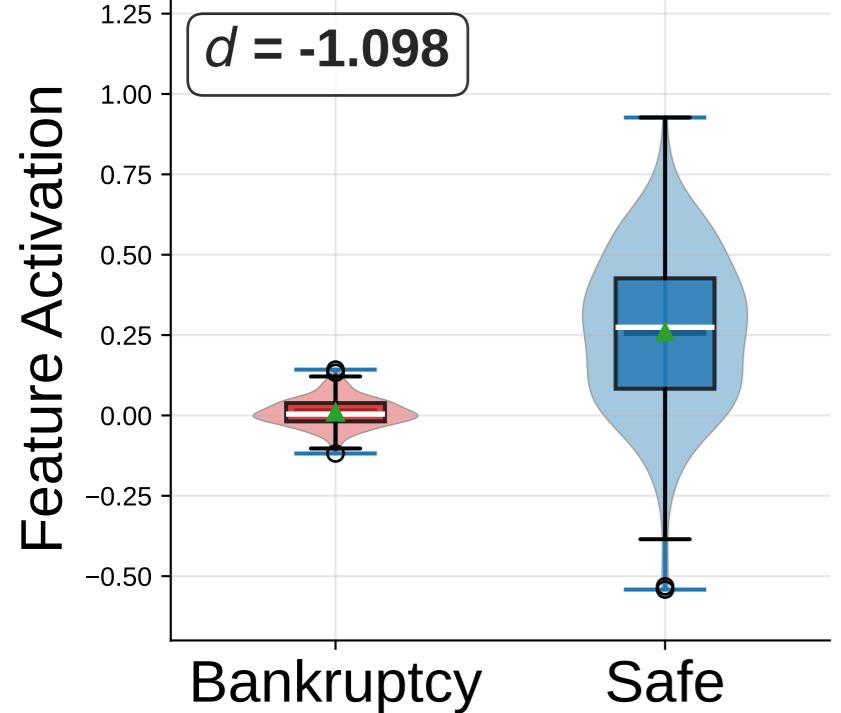
Layer 26 - Feature 31257



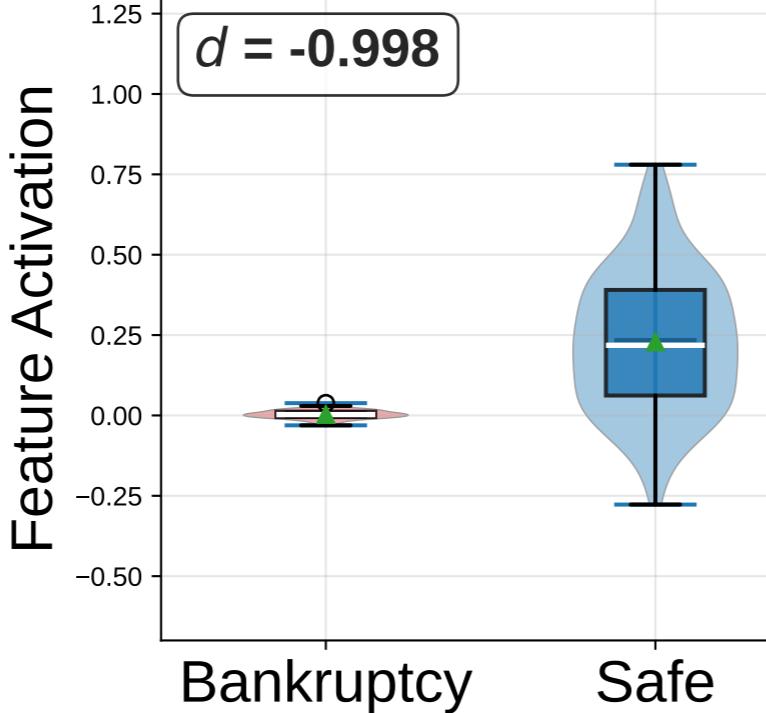
Layer 27 - Feature 32051



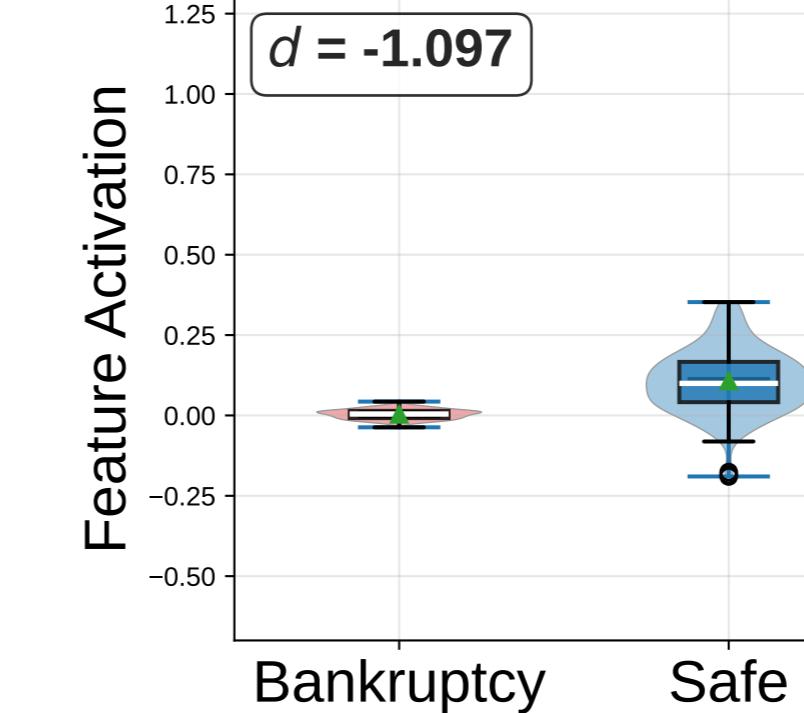
Layer 28 - Feature 21135



Layer 29 - Feature 29399



Layer 30 - Feature 17910



Layer 31 - Feature 29127

