"is_fake" Distribution 0.5 -0.4 proportion 5.0 0.2 data_type 0.1 test train val 0.0 0 is_fake