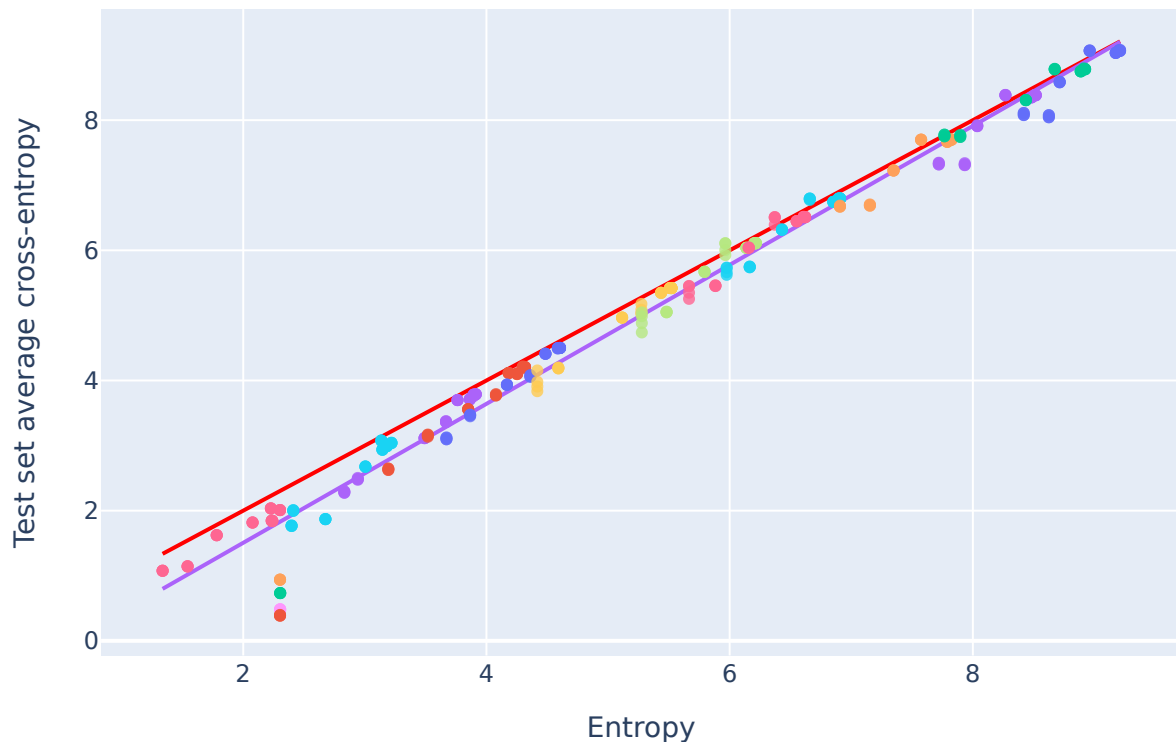


Test set average cross-entropy vs. entropy, $R^2=0.9842$



Vocab size

— Theoretical limit	— $1.07 * x + -0.63$	● 10	● 25	● 26	● 50
● 60	● 75	● 100	● 250	● 283	● 500
● 2500	● 5000	● 7500	● 8193	● 10000	