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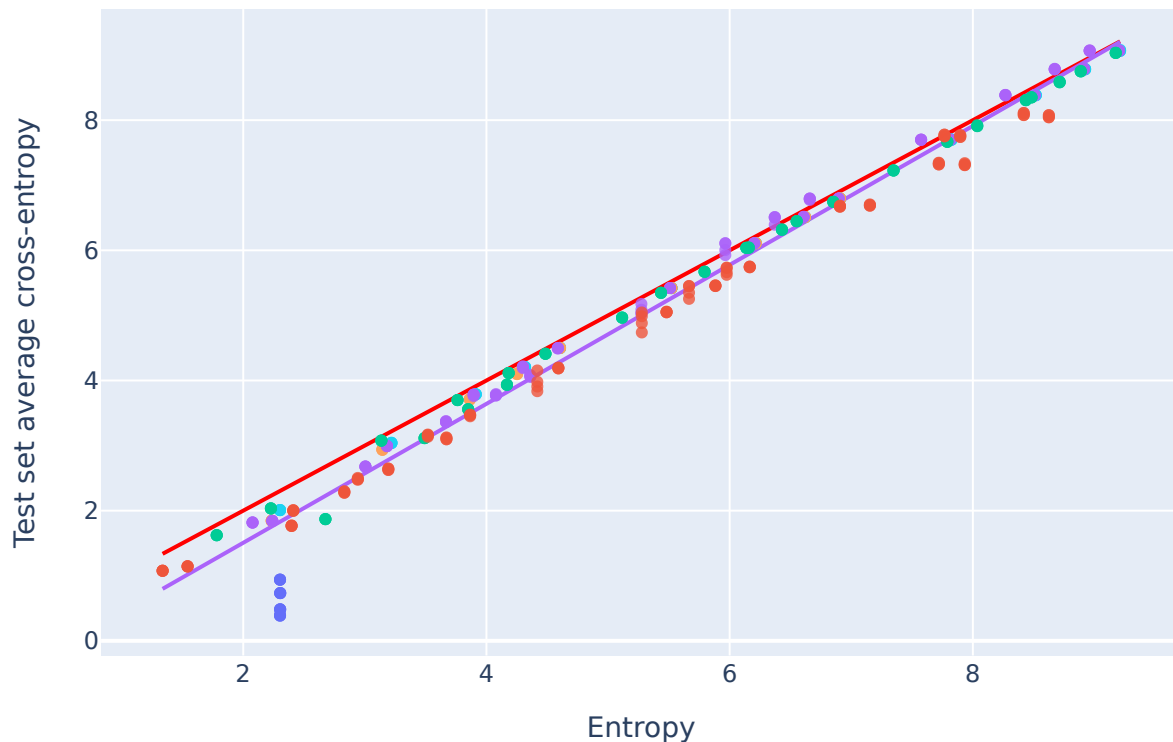
Test set average cross-entropy

Entropy

Uni- or bigram and distribution

— Theoretical limit — 1.07 * x + -0.63 • unigrams, uniform
• bigrams, uneven • bigrams, normal • unigrams, normal
• bigrams, long_range • unigrams, manual