Mistral-7B on the *ASAP* dataset

Pearson's r: -0.43; p: 10^{-93}

average entropy per token

essay grade (%)