

Correlation of BERTScore to human judgement
for pairs with some lexical overlap

