

PCA projection of hidden state of
Dirty RoBERTa-Sentinel on OpenGPTText-Final

