Robert Wolfe
November 29, 2022
Erratum to 2021.emnlp-main.41: Low Frequency Names Exhibit Bias and Overfitting in
Contextualizing Language Models

We have identified a bug in our code which affects the inter-layer similarity measurement
based on the linear CKA method. Specifically, the matrix of representations from each layer was
errantly transposed prior to being input to the linear CKA function. After the transpose is
corrected, we could not show that statistically significant correlations exist between name
frequency and inter-layer similarity in the contextualizing language models studied.