Top 5 Retained 0.00 0.36 0.37 0.45 0.55 0.47 0.53 0.61 0.62 0.60 0.59 0.59 0.59

BERT - DuoRC Integrated Gradients JSD

0.36 0.00 0.35 0.39 0.47 0.46 0.47 0.55 0.59 0.59 0.58 0.58 0.58

0.37 0.35 0.00 0.31 0.40 0.40 0.40 0.47 0.48 0.48 0.46 0.46 0.44

0.45 0.39 0.31 0.00 0.33 0.37 0.34 0.40 0.44 0.45 0.45 0.47 0.45 0.55 0.47 0.40 0.33 0.00 0.40 0.32 0.39 0.48 0.52 0.50 0.50 0.49

0.47 0.46 0.40 0.37 0.40 0.00 0.30 0.36 0.39 0.39 0.46 0.48 0.47

0.53 0.47 0.40 0.34 0.32 0.30 0.00 0.26 0.32 0.36 0.41 0.44 0.42

0.61 0.55 0.47 0.40 0.39 0.36 0.26 0.00 0.18 0.24 0.37 0.42 0.38

0.62 0.59 0.48 0.44 0.48 0.39 0.32 0.18 0.00 0.12 0.24 0.31 0.26

0.60 0.59 0.48 0.45 0.52 0.39 0.36 0.24 0.12 0.00 0.22 0.28 0.24

0.59 0.58 0.46 0.45 0.50 0.46 0.41 0.37 0.24 0.22 0.00 0.12 0.08

0.59 0.58 0.46 0.47 0.50 0.48 0.44 0.42 0.31 0.28 0.12 0.00 0.06

0.59 0.58 0.44 0.45 0.49 0.47 0.42 0.38 0.26 0.24 0.08 0.06 0.00

10