

Detecting Persuasive Atypicality by Modeling Contextual Compatibility (Supplementary File)

Meiqi Guo Rebecca Hwa Adriana Kovashka
Department of Computer Science, University of Pittsburgh, Pittsburgh PA, USA
meiqi.guo@pitt.edu {hwa, kovashka}@cs.pitt.edu

A. Significance tests on our models

For showing the significance of improvement by our unsupervised models over baselines, we report the standard error of the best AUCs and p-value (the probability of a non-best AUC equals best, given samples) in Table 1.

Model	TR1	TR2	OIO	OMP	CP
AE	54.67 <small>p=.03</small>	63.28 <small>p=.41</small>	38.79 <small>p=.00</small>	52.98 <small>p=.09</small>	57.78 <small>p=.02</small>
SVM	64.81 <small>p=.50</small>	68.27 <small>p=.98</small>	59.36 <small>p=.17</small>	65.81 ± 5.5	54.21 <small>p=.00</small>
T	62.66 <small>p=.09</small>	60.72 <small>p=.03</small>	63.07 <small>p=.16</small>	42.52 <small>p=.00</small>	69.18* <small>p=.54</small>
RST	67.50 ± 3.8	68.37 ± 4.1	67.31 ± 5.4	55.18 <small>p=.03</small>	71.26* ± 3.7
Model	SDO	LDO	OR	Others	MIC AVE
AE	56.62 <small>p=.05</small>	56.05 <small>p=.23</small>	48.57 <small>p=.00</small>	50.99 <small>p=.01</small>	52.52 <small>p=.00</small>
SVM	65.12 <small>p=.40</small>	54.43 <small>p=.07</small>	56.31 <small>p=.04</small>	54.23 <small>p=.04</small>	58.82 <small>p=.02</small>
T	63.71 <small>p=.22</small>	61.63 <small>p=.53</small>	64.05* ± 2.9	63.68* ± 3.4	62.86* <small>p=.36</small>
RST	68.67 ± 4.8	63.99 ± 4.4	61.84 <small>p=.37</small>	59.68 <small>p=.14</small>	64.32* ± 2.0

Table 1. Unsupervised models on Ads; best AUC with its standard error; p-value with respect to the best one for others; our model (RST/T) is starred if significantly better than baselines ($p < .1$).