Appendix E. Monte-Carlo dropout uncertainty estimates of BERTs.

Classifier	Label	Train Proportion	Mean	SD	Variance
1A	No	11.2	0.170	0.326	0.0024
1A	Yes	88.8	0.830	0.326	0.0024
1B	Not applicable	88.8	0.861	0.321	0.0011
1B	No	2.2	0.018	0.095	0.0001
1B	Yes	9.0	0.120	0.300	0.0010
2A	Not applicable	2.0	0.026	0.111	0.0002
2A	No	67.5	0.694	0.426	0.0010
2A	Yes	30.5	0.280	0.418	0.0008
2B	Not applicable	2.0	0.020	0.087	0.0003
2B	No	88.5	0.874	0.230	0.0011
2B	Yes	9.5	0.106	0.201	0.0007
2C	0	6.8	0.083	0.190	0.0005
2C	1	5.2	0.058	0.158	0.0002
2C	2	33.8	0.332	0.429	0.0006
2C	Not applicable	2.0	0.023	0.042	0.0001
2C	Undefined	51.7	0.503	0.467	0.0010
3A	Not applicable	2.0	0.021	0.121	0.0006
3A	No	93.0	0.920	0.258	0.0007
3A	Yes	5.0	0.059	0.221	0.0001
4A	Not applicable	2.0	0.010	0.031	0.0000
4A	No	41.8	0.456	0.457	0.0018
4A	Opinion	18.0	0.170	0.350	0.0010
4A	Yes	38.2	0.357	0.435	0.0019

Note. Each row shows uncertainty statistics from Monte Carlo dropout predictions for a BERT classifier-label pair. The “Train Proportion” refers to the percentage of training samples for that label. “Mean” and “SD” are the average and standard deviation of prediction confidence across utterances. “Variance” is the mean of the squared SDs across utterances, capturing overall uncertainty more accurately.