A Supplemental Material

A.1 Model Performances in the Original Facebook Dataset (Without Anonymization)

Model performances in the main paper were based on the anonymized dataset (see 1). Here, we show results on the original, non-anonymized data (i.e. without named entities and dates removed). These models are trained with the orignal dataset and used for the pipeline application for the exploration in Facebook (Kosinski et al., 2013) and Yelp review dataset (Zhang et al., 2015).

A.2 Performances on Validation Dataset

We also report the validation F1s of the optimized model for each subtask performance of the complete pipeline: the feature-based models in causality prediction (CP) and the LSTM variants in causal explanation identification (CEI). We saw the same pattern of results comparing LSTMs but the ablation analyses were not as conclusive. We note that our hyper-parameters were optimized over the validation set and therefore these results, as opposed to those in the main paper, might be slightly overfit.

References

Or Biran and Kathleen McKeown. 2015. Pdtb discourse parsing as a tagging task: The two taggers approach. In *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 96–104.

Michal Kosinski, David Stillwell, and Thore Graepel. 2013. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15):5802–5805.

Ziheng Lin, Hwee Tou Ng, and Min-Yen Kan. 2014. A pdtb-styled end-to-end discourse parser. *Natural Language Engineering*, 20(02):151–184.

Xiang Zhang, Junbo Zhao, and Yann LeCun. 2015. Character-level convolutional networks for text classification. In *Advances in neural information processing systems*, pages 649–657.

Tags	Contents
ANONYMIZED_ALPHANUMERIC	phone numbers with '-', user names on websites, etc
ANONYMIZED_AMOUNT	specific amount of money or other quantities
ANONYMIZED_DATE	specific dates for anniversaries, etc: month, day, year
ANONYMIZED_LOCATION	specific locations for travels, hometown, etc
ANONYMIZED_NAME	names of people, streets, buildings, etc
ANONYMIZED_NUMBER	street numbers of addresses, phone numbers, etc
ANONYMIZED_TIME	specific time for events such as meetings, parties, etc
ANONYMIZED_URL	URLs for personal websites, games, etc

Table 1: Anonymization tags which were used in our anonymized dataset

Causality Messages	CE DA	Total DA
Training	1,278	5,609
Validation	160	653
Test	160	759
Total	1,598	7,021

Table 2: The number of discourse arguments in the original causality messages.

Model	F1
(Biran and McKeown, 2015)	0.434
(Lin et al., 2014)	0.640
Linear SVM	**0.788**
RBF SVM	0.743
Random Forest	0.783
LSTM	0.743

Table 3: Causality prediction performance across different predictive models. Bold indicates significant improvement over the LSTM.

Model	F1
All	0.788
- First-Last, First3	0.790
- Word Pairs	0.806
- POS tags	0.746
- (Char + Word) N-grams	0.756
- Sentiment tags	0.788

Table 4: Feature ablation test of Linear SVM for causality prediction

Model	Prec	Rec	F1
Linear SVM	0.767	0.733	0.746
RBF SVM	0.756	0.771	0.762
Random Forest	0.754	0.793	0.752
LSTM	0.859	0.864	0.861

Table 5: Causal explanation identification performance.

Model	CP (F1)	CEI (F1)
Full LSTM	0.743	0.861
DA AVG LSTM	0.738	0.808
Word LSTM	0.655	0.802

Table 6: The effect of Word-level LSTM for causality prediction (CP) and causal explanation identification (CEI) in the original CEI test set.

Model	Prec	Rec	F1
CP + CEI_{causal}	**0.870**	**0.883**	**0.873**
CP + CEI_{all}	0.854	0.875	0.856
CEI_{causal} Only	0.851	0.795	0.815
CEI_{all} Only	0.842	0.853	0.847

Table 7: The effect of DA-Level LSTM for causal explanation identification. Bold: significant ($p < .05$) increase in F1 over the next best model.

Model	A F1	O F1
All	0.760	0.767
- First-Last, First3	0.765	0.767
- Word Pairs	0.764	0.757
- POS tags	0.746	0.723
- (Char + Word) N-grams	0.778	0.776
- Sentiment tags	0.757	0.767

Table 8: Feature ablation test of Linear SVM for the causality prediction (CP) on the validation sets of the anonymized dataset (A) and the original dataset (O)

Model	A F1	O F1
Full LSTM	0.846	0.851
DA AVG LSTM	0.813	0.815
Word LSTM	0.765	0.766

Table 9: The validation F1s of the architectural variants of the causal explanation identification (CEI) LSTMs.