A Time Ranges for Time Difference Embeddings

We encode the ranges of timestamp difference between the predicted message m_i and an arbitrary message m_j in the past, ie. TimeDiff(m_i, m_j) = TimeStamp(m_i) − TimeStamp(m_j) into 11 buckets, shown in Table 6. Then, each bucket is represented by a 20-dimensional embedding in Reddit experiments

TimeDiff(m_i, m_j)	index
if m_j is PAD	0
0 min.	1
(0mins, 5mins]	2
(5mins, 30mins]	3
(30mins, 1hours]	4
(1hours, 3hours]	5
(3hours, 6hours]	6
(6hours, 12hours]	7
(12hours, 24hours]	8
(24hours, 72hours]	9
>72hours	10

Table 6: Ranges of timestamp difference

B Reddit Dataset Statistics and Examples

In Table 7, we shows the detailed statistics of the three Reddit datasets, including total numbers of threads and messages, their counterparts for train/validation/test sets and message size per thread. Finally, each Reddit comment has a message id to which it replies in the original dataset. Due to the pre-processing steps in section 3, not all the original comments are kept after the data construction procedure. Therefore, we record the percentage of messages whose "reply-to" messages are still present in the dataset. Note that, following (Jiang et al., 2018), this information is not used in the experiments, because a typical chat channel can not have an explicit indication of the "reply-to" parent for every message.

Dataset	Gadgets	Iphones	Politics
train threads	230	539	2950
train messages	7.4k	11.6k	84.8k
valid threads	30	67	368
valid messages	918	1.6k	10.2k
test threads	28	68	370
test messages	1k	1.4k	9.9k
msg. per thread	32.6	21.7	28.4
msg. w/ "reply-to"	73.9%	71.9%	76.2%

Table 7: Dataset statistics

In Table 8, we show a chunk of the real message sequence in Politics dataset. They are sorted by time stamps. For each message, there are 6 attributes. They are thread id, message id, time stamp (in seconds), message body, username, respectively. For example, the messages in blue (1st, 3rd and 4th) belong to the same thread, whose starting message is the first one in the table.

thread id	message id	time stamp	message body	username
. . .	. . .	. . .	. . .	. . .
t3_4lyd8p	t3_4lyd8p	1464739232	Cubii Desk Elliptical: Might As Well Work Out While I Work, Right?	tembbo
t3_4lxskc	t1_d3r5xnp	1464743587	I was sold until it said it comes with Windows 10 pre-installed.	B4U
t3_4lyd8p	t1_d3r67eo	1464744032	I have used both a treadmill desk and a bike desk. Treadmill desk is fine for mail and writing/using Word. Powerpoint and Excel which require fine hand movements, I still prefer a normal desk. Not very good with the bike desk...can basically just browse. On the treadmill desk, my "bottom half" goes on autopilot.	tembbo
t3_4lyd8p	t1_d3r6byp	1464744238	Maybe different stokes for different folks? I suck at biking - can relate about not being able to work on a bike desk.	phhhhart
t3_4loba7	t1_d3r6klt	1464744620	People seem to make this mistake a lot., It's also worth noting that if you've been into subcompacts, ultraportables, netbooks, etc, you're probably used to putting up with some dongles or breakouts or other inconveniences due to the form factor.	Batman
. . .	. . .	. . .	. . .	. . .

Table 8: Data example in Politics dataset