

# Numeracy-600K: Experimental Results Update - 20190824

Chung-Chi Chen

cjchen@nlg.csie.ntu.edu.tw

## 1 Experimental Results Update

Because I made a mistake during preprocessing, I rebuilt the dataset and rerun the experiments. The following tables list the latest results. The distributions of each set are slightly different from those of the previous version. BiGRU still performs the best in both sets. Most of models get a little bit improvement.

I would like to thank Chengyue Jiang from ShanghaiTech University for pointing out the error.

Magnitude	Range	Ratio
Decimal	$0 \leq m < 1$	24.14
1	$1 \leq m < 10$	38.50
2	$10 \leq m < 10^2$	24.04
3	$10^2 \leq m < 10^3$	11.72
4	$10^3 \leq m < 10^4$	1.07
5	$10^4 \leq m < 10^5$	0.29
6	$10^5 \leq m < 10^6$	0.23
> 6	$10^6 \leq m$	0.01

Table 1: Distribution of numerals in the dataset.

Model	Micro-F1	Macro-F1
LR	72.78%	61.83%
CNN	78.42%	60.45%
GRU	79.65%	60.72%
BiGRU	<b>81.04%</b>	<b>65.94%</b>
CRNN	79.29%	65.78%
CNN-capsule	77.20%	63.60%
GRU-capsule	79.55%	65.79%
BiGRU-capsule	78.91%	65.36%

Table 2: Experimental results of comments.

M	0	1	2	3	4	5	6	7
%	0.31	36.38	31.47	6.88	24.30	0.34	0.31	0.01

Table 3: Distribution of numerals in the title dataset. M.: magnitude; 7:  $M > 6$ .

Model	Micro-F1	Macro-F1
LR	63.86%	36.91%
CNN	69.12%	38.81%
GRU	71.83%	42.92%
BiGRU	71.87%	<b>43.96%</b>
CRNN	68.84%	40.22%
CNN-capsule	63.84%	30.12%
GRU-capsule	<b>72.03%</b>	34.28%
BiGRU-capsule	71.86%	34.15%

Table 4: Experimental results of titles.

Training	Test set	Micro-F1	Macro-F1
Comment	Title	31.68%	10.80%
Title	Comment	26.25%	10.75%

Table 5: Results of learning cross-source numeracy.