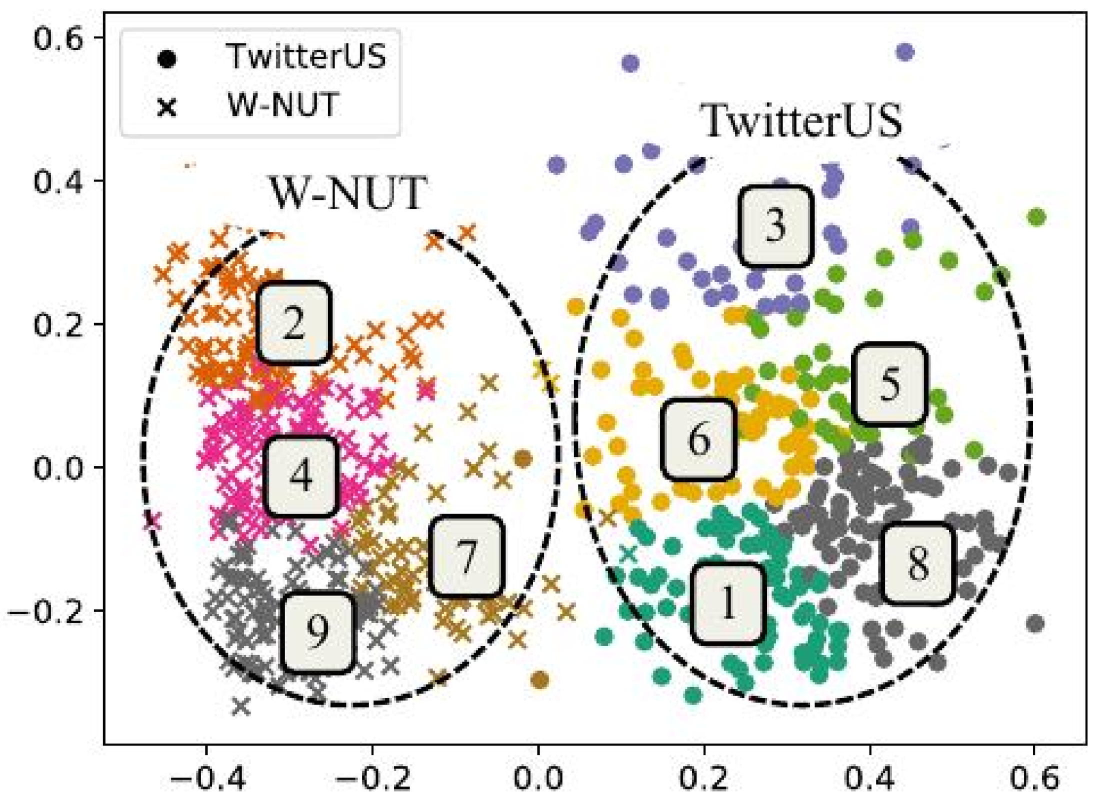

Cluster ID	Dataset	Timeline	Location	Description	Timezone	User	User Network
1	TwitterUS	0.843	0.082	0.040	0.035	0.359	0.641
2	W-NUT	0.517	0.317	0.081	0.085	0.732	0.268
3	TwitterUS	0.432	0.430	0.069	0.069	0.319	0.681
4	W-NUT	0.637	0.160	0.097	0.105	0.737	0.263
5	TwitterUS	0.593	0.219	0.114	0.075	0.230	0.770
6	TwitterUS	0.672	0.214	0.069	0.045	0.365	0.635
7	W-NUT	0.741	0.077	0.080	0.102	0.605	0.395
8	TwitterUS	0.766	0.099	0.068	0.067	0.222	0.778
9	W-NUT	0.800	0.067	0.056	0.078	0.730	0.270

Figure 6: A k-means clustering result and the attention probabilities of users that are closest to the cluster centroids. The underlined values are the max values of the two datasets for each column.

Model	Error Distance		
	Median	Mean	σ
Oracle	23.3	31.4	30.1

Table 4: Error distance values in TwitterUS with oracle predictions. σ in the table denotes the standard deviation.

measured the performance of an oracle model where city predictions are all correct (accuracy of 100%) in the test set.

Table 4 denotes this oracle performance. The oracle mean error distance is 31.4 km. Its standard deviation is 30.1. Note that ground truth locations of TwitterUS are geotags and will not exactly match the oracle city centers. These oracle values imply that the current median error distances are close to the lower bound of the city classification approach and that they are difficult to improve.

6.2.2 Errors with High Confidences

The proposed model still contains 28–30% errors even in accuracy@161. A qualitative analysis of errors with high confidences was performed to investigate cases that the model fails. We found two common types of error in the error analysis. The first is a case when a location field is incorrect due to a reason such as a house move. For example, the model predicted “Hong Kong” for a user with a location field of “Hong Kong” but has the gold location of “Toronto”. The second is a case when a user tweets a place name of a travel. For example, the model predicted “San Francisco” for a user who tweeted about a travel to “San Francisco” but has the gold location of “Boston”.

These two types of error are difficult to handle with the current architecture of the proposed model. The architecture only supports single location field which disables the model to track location changes. The architecture also treats each tweet independently which forbids the model to express a temporal state like traveling.

7 Conclusion

As described in this paper, we proposed a complex neural network model for geolocation prediction. The model unifies text, metadata, and user network information. The model achieved the maximum of a 3.8% increase in accuracy and a maximum of 6.6% increase in accuracy@161 against several previous state-of-the-art models. We further analyzed the states of several attention layers, which revealed that the probabilities assigned to timeline representations and user network representations match to some statistical characteristics of datasets.

As future works of this study, we are planning to expand the proposed model to handle multiple locations and a temporal state to capture location changes and states like traveling. Additionally, we plan to apply the proposed model to other social media analyses such as gender analysis and age analysis. In these analyses, metadata like location fields and timezones may not be effective like in geolocation prediction. However, a user network is known to include various user attributes information including gender and age (McPherson et al., 2001) which suggests the unification of text and user network information to result in a success as in geolocation prediction.

Acknowledgments

We would like to thank the members of Okumura–Takamura Group at Tokyo Institute of Technology for having insightful discussions about user profiling models in social media. We would also like to thank the anonymous reviewer for their comments to improve this paper.