

Figure 1: The distribution of errors in the U.S. states over the development set of TWITTER-US using three geolocation models: a) GCN, b) MLP-TXT+NET, and c) DCCA using only 1% of samples as labelled data. The diameter of the circles shows the number of development users in each state, and the colour indicates Median error in km.

(c) DCCA

MS AL

GA

FL

240

160

LA

ΤХ

AZ



Supplementary material for Rahimi et al. (2018). Semi-Supervised User Geolocation via Graph Convolutional Networks (ACL2018).

Figure 2: Confusion matrix between the known and predicted states of development users of TWITTER-US using the GCN geolocation model based on only 1% of samples as labelled data.