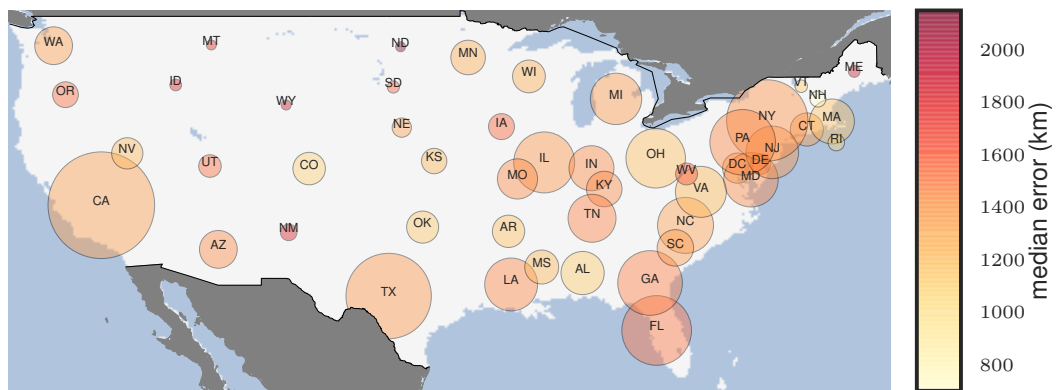
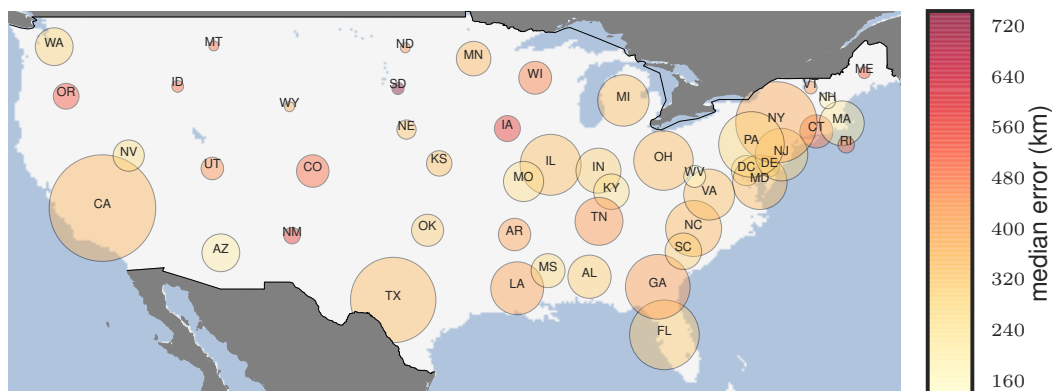


(a) GCN



(b) MLP-TXT+NET



(c) DCCA

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.

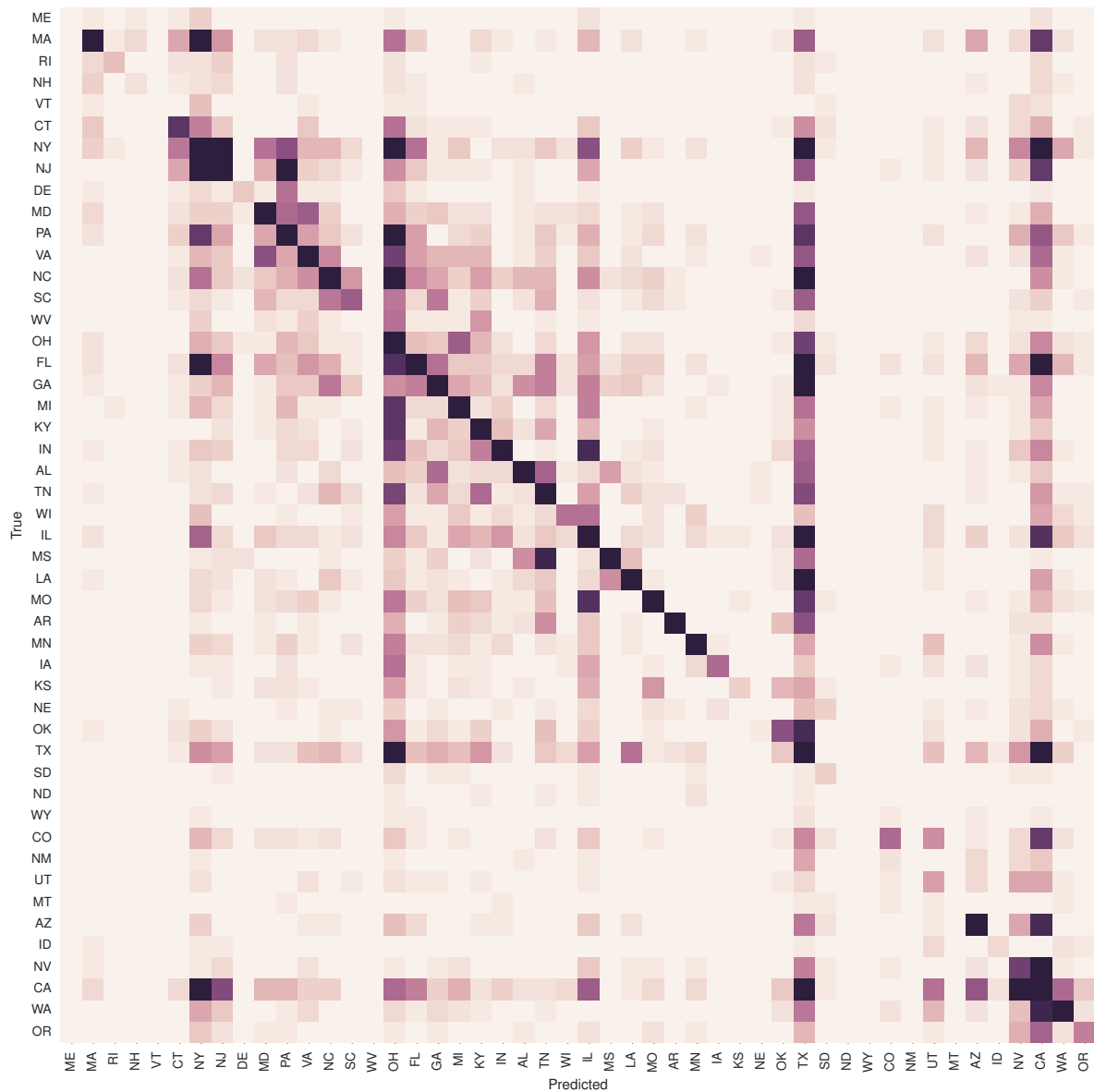


Figure 2: Confusion matrix between the known and predicted states of development users of TWITTER-US using the GCN geolocation model based on only 1% of samples as labelled data.