Checkpoint 5: Natural Language Processing

COMP_SCI 396/496: Data Science Seminar
Meenakshi Kommineni, Madison McClellan, Archana Ramasubramaniam

Do civilian narratives of complaint reports against officers in a high percentile of TRRs (90th percentile and above) contain more negative sentiment than those against officers in a low percentile of TRRs (below 50th percentile)?

To answer this question, we used the built-in sentiment analyzer in Python's Natural Language Toolkit (NLTK). We gathered data by connecting to the PostgreSQL database from within the Jupyter Notebook and querying the desired data. In our first attempt, we decided to partition the data into the two percentile groups prior to training the model. The first dataset consisted of complaint reports (CRs) against officers in the 90th percentile or above of TRRs and the second consisted of CRs against officers in the 50th percentile or below. We then ran the NLTK model over both datasets to calculate the number of CRs with positive and negative sentiment. We then calculated the proportion of positive and negative sentiment for both datasets. For the >=90th percentile dataset, the proportion of CRs with positive sentiment was 10.60% and the proportion with negative sentiment was 88.78%. For the <=50th percentile dataset, the proportion of CRs with positive sentiment was 11.0% and the proportion with negative sentiment was 88.51%. In our second attempt, we queried all of the CRs and trained the model over the whole dataset and partitioned the data into strata afterwards. For CRs against officers in the >=90th percentile group, 10.60% included positive sentiment and 88.54% included negative sentiment. For CRs against officers in the <=50th percentile group, 11.51% included positive sentiment and 88.0% included negative sentiment. These results are summarized in the table below.

	Attempt #1 (partition data before training)	**Attempt #2 (train data over all CRs)**
>= 90th percentile	Positive: 10.60% Negative: 88.78%	Positive: 10.60% Negative: 88.54%
<= 50th percentile	Positive: 11.0% Negative: 88.51%	Positive: 11.51% Negative: 88.0%

The first takeaway from these results is that both approaches achieved almost identical results. The second takeaway is that there is a very minor difference (almost none) between sentiment expressed towards officers in the different groups. Initially we expected that CRs against officers in the >=90th percentile group would have much more negative sentiment than those against officers in the <=50th percentile group. However, the data disproves this. This is an interesting finding because it might suggest that behavioral training of officers in the high TRR percentile group will not make a big difference because they are already perceived the same as the officers in the lower TRR percentile group. Rather, increasing the severity of punishment for unwarranted uses of force might be a more effective approach.

Finally, we wanted to evaluate how accurate the NLTK models are against the CR narratives. To do this, we gathered a small sample of CRs and examined how the model classified them. The narratives that were classified as negative contained stronger language, mentioned weapons, and included words that are generally perceived as negative. However, the narratives that were classified as positive did not contain many words that are generally perceived as positive. Rather, these narratives were “less negative” than those classified as negative. For example, they did not mention weapons and typically involved less severe incidents such as traffic stops.