

1. Switch to conditional probability for linguistic features:

Recall $p(X_{sent} = x_i) = \frac{P(x_i) \in C'_i}{\sum_{i=1}^z f(x_i) \in C'_i}$ is equivalent to the number of comments (tweets) with sentiment x_i divided by the total number of comments. Therefore, $p(X_{caps} = x_i) = p(X_{sent} = x_i | C'_i \in Caps) = \frac{p(X_{sent}=x_i \cap C'_i \in Caps)}{p(C'_i \in Caps)}$ where $Caps$ is the set of comments that have capitalized terms. This works similarly for the extreme lexicon.