

Task 5 - Creative Exploration

So far you have used basic models to understand and predict words. In this next task, your goal is to use all the resources you have available to you (from the Data Science Specialization, resources on the web, or your own creativity) to improve the predictive accuracy while reducing computational runtime and model complexity (if you can). Be sure to hold out a test set to evaluate the new, more creative models you are building.

Tasks to accomplish

1. Explore new models and data to improve your predictive model.
2. Evaluate your new predictions on both accuracy and efficiency.

Questions to consider

1. What are some alternative data sets you could consider using?
2. What are ways in which the n-gram model may be inefficient?
3. What are the most commonly missed n-grams? Can you think of a reason why they would be missed and fix that?
4. What are some other things that other people have tried to improve their model?
5. Can you estimate how uncertain you are about the words you are predicting?