

NEST Labs Bootcamp, August 2017

Day 5

“The best programmers are not marginally better than merely good ones. They are an order-of-magnitude better, measured by whatever standard: conceptual creativity, speed, ingenuity of design, or problem-solving ability.” – Randall E. Stross.

Code fast, code smart, and git your codes running!

1. Okay, here's the deal. We need to split up an entire corpus into meaningful sentences. While this seems to be a very simple task; perhaps by writing a regular expression that splits by “a punctuation followed by a space then a capital letter” e.g. How are you? Hope you're good. Splits by question mark (punctuation) followed by a space, then a capital letter “H” (Hope), we get [“How are you?”, “Hope you're good.”].

But after reviewing this method across multiple paragraphs we ran into an issue. For example, What if we had a sentence that goes like this... “Mr. Foo went to the bar”. Splits by a period (.) i.e. a punctuation followed by a space then capital letter “F”(Foo), we get [“Mr.”, “Foo went to the bar”]. However, that is not a complete sentence.

Your task is to implement a method by which we can split a body of text into meaningful sentences.

2. The lead programmer for Google Hangouts after merging all the project's source code wondered how many lines of codes are there in the project? However, he used so many third party libraries, which doesn't count as part of the code he and his team wrote.

Therefore, he contacted us (NESTLabs) to solve this problem for him. His specs were to uncheck all third party library's folders & files. However this behavior can be

simulated in whichever way feels convenient for you to implement.

3. Implement any sorting algorithm of your choice or more interestingly, you can come up with your own.

God speed!