

Programming Assignment 2

1 Assignments

1. Change your MapReduce program to BigramCount.java
2. Change the contents in “input” folder as follows
 - file0: contains ~100 lines of text
 - file1: contains ~150 lines of text
 - file2: contains ~50 lines of text
3. Modify your mapper so that it carries out the following
 - Count bigrams (pairs of consecutive words) instead of individual words; delimit words with space
 - If a line contains only one word, skip this line
 - Figure out how many `map` tasks are launched in the `map` stage
 - Do not use built-in counters

2 Submission

- In `answer.txt`, (1) explain how to count bigrams and figure out how many `map` tasks are launched in the `map` stage, and (2) provide your solution to the number of `map` tasks
- `tar -cvf pa2_<your last name>.tar input output BigramCount.java answer.txt`
- Upload the tar file to blackboard before the deadline
 - Deadline: February 5 (Tuesday), 2019 @ 11:59AM