

Module 3: Hadoop MapReduce Framework

Assignment - Word Count

edureka!

edureka!

© 2014 Brain4ce Education Solutions Pvt. Ltd.

Module 3: Hadoop MapReduce Framework

Assignment – Count the number of occurrences of each word in a text file

Table of Contents

Introduction	2
Problem statement	2

edureka!

Introduction

Apply your MapReduce programming knowledge and write a MapReduce program to process a text file. You need to print the count of number of occurrences of each word in the text file.

The dataset for this problem is the text file '**wordcountproblem**' available in your LMS

Problem statement

Let's understand the problem through a sample text file content:

"Hello everyone this is a sample dataset. You need to print the word count of particular words in this dataset."

Your MapReduce program should process this text file and should provide output as follows:

Output

Word	Word Count
------	------------

a	1 (As the word 'a' occurred only once)
this	2 (As the word 'this' occurred twice)

Your task in this assignment is to process the '**wordcountproblem**' text file using MapReduce program and count the number of occurrences of each word in this text file .

You can review the solution in your LMS.