Vidyavardhini's College of Engineering and Technology
Department of Artificial Intelligence & Data Science

AY: 2024-25

Class:	BE	Semester:	VII
Course Code:		Course Name:	Big Data Analytics

Name of Student:	Hitesh. A. Moota
Roll No. :	32
Assignment No.:	OL
Title of Assignment:	
Date of Submission:	
Date of Correction:	

Evaluation

Performance Indicator	Max. Marks	Marks Obtained
Completeness	5	3
Demonstrated Knowledge	3	2
Legibility	2	1
Total	10	6

Performance Indicator	Exceed Expectations (EE)	Meet Expectations (ME)	Below Expectations (BE)
Completeness	5	3-4	1-2
Demonstrated Knowledge Legibility	3	2	1
Legibility	2	1	0

Checked by

Name of Faculty :

Signature :

Date :

BDA

Assignment -2

Q.1) Solve following problem using MapReduce word count method " The doctor ate an apple. The Engineer ate mango & banana". Also write the pseudo code or word count for by using MapReduce algorithm.

→ Pseudo code or word count by using MapReduce:-

Word count using MapReduce
→ Count occurrence of each word input.
M = large corpus of text
Mapper:-
    Each word in $M \rightarrow$ ("word", 1)
Reducer:-
    Each word in $M \rightarrow$ ("word", 0 ?)
        $\rightarrow$ ("word", Sum is ?)
    " aggregate in Hadoop.

○ Job :- Count one Occurrences of each word in a data set

○ Sentence:- The doctor ate an apple. The Engineer ate mango & banana.

MapReduce word count diagram

Input:
- "the doctor ate an apple"
- "The Engineer ate mango & banana"

Map tasks — Map (key, value) pair splitting:
- the,1 / doctor,1 / ate,1 / an,1 / apple,1
- the,1 / Engineer,1 / ate,1 / mango,1 / and,1 / banana,1

Sort & Shuffle:
the,1 / the,1 / doctor,1 / ate,1 / ate,1 / an,1 / apple,1 / Engineer,1 / mango,1 / and,1 / banana,1

Reduce tasks — Reduce (key, value) pair:
the,2 / doctor,1 / ate,2 / an,1 / apple,1 / Engineer,1 / mango,1 / and,1 / banana,1

Final Output:
the,2 / doctor,1 / ate,2 / an,1 / apple,1 / Engineer,1 / mango,1 / and,1 / banana,1

FOR EDUCATIONAL USE

.) write pseudo code for union by mapreduce & Explain
it with the help of example

→ Pseudo code for Union :-

Map (key, value):
  - for tuple in value:
      unit (tupple, tupple).

reduce (key, values):
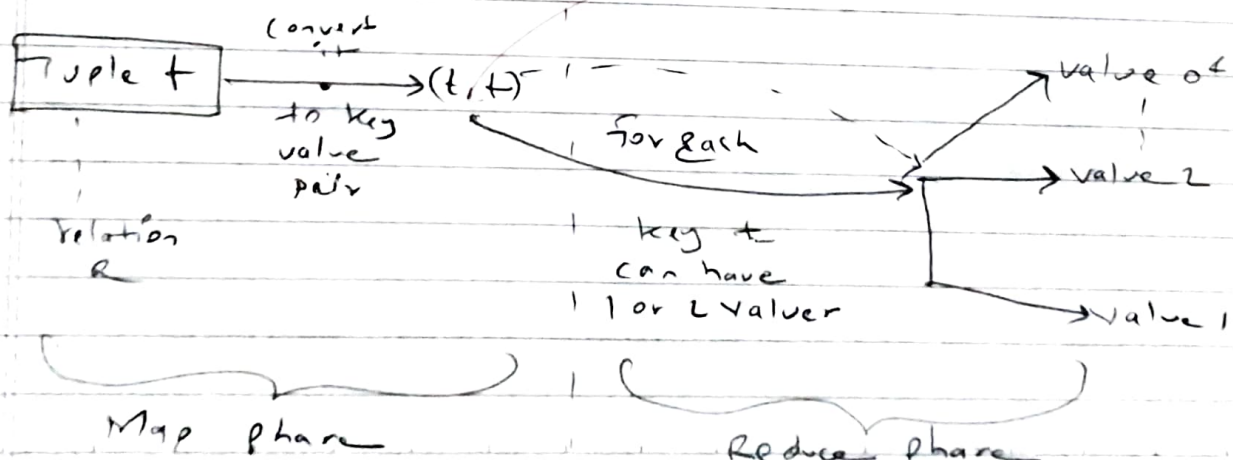 Emit (key, key)

In other words

Class Mapper
    method Map (revokey key, tuple t)
    Emit (tuple t, null)

Clasr Reduce
    method Reduce (tuple t, arrayn) // n is an array
of one //
    Emit (tuple t, null)   or two nulls

| Tuple t | Convert it →(t,t) | | value of |

to key
value
pair

relation
R

for each

key t
can have
1 or 2 valuer

value 2

value 1

Map phase   Reduce phase

Example :-

Union : $R_1 \cup R_2$

SELECT * FROM R_1

UNION

SELECT * FROM R_2

R_1

A	B
a_1	b_1
a_2	b_1

$\cup$

R_2

A	B
a_1	b_1
a_3	b_4

=

$R_1 \cup R_2$

A	B
a_1	b_1
a_2	b_1
a_3	b_4

Name	Age
John	21
Tom	17
Mike	22
Smith	16

Name	Age
Monty	21
Tom	19
Dean	22
Smith	16

Union →

Name	Age
John	21
Tom	17
Mike	22
Smith	16
Monty	21
Tom	19
Dean	22