PO	OPU°
Topi	Date 8/3/4041
ТОР	Page No
	Experiment-8
	Ain: Word Count problem and caching
0	Wate a sport by oan a single single
1	Read file spark. +xt
	text-pile = sc. textfile ("file Store / tables/spack. +x+")
2	Split the file
$\rightarrow$	doita split = text file. flattap (lambda 2:x.speit(" )
	は素練・生き合い コー・コー・ファイン さいさいき きこう
3	Make a fair RDD by adding 1 to each word.  data pair = data-split, map (lambda word:
3	data brier = data split man (landed a survey)
	( word, 1))
	(0000, 1)
4	word court
1	
	courts = data-pair. reduce By key (Cambda ab:
2	sauce the result întopèle.
	counts, saneAsTextFile ("dbfs: /FileStore/tables/resultst
-	
6	Verify the result.
5	sc. textfile (" albs: / Filstone / tables/result. +x4"). rake (5)
2	Caching
1	
1	context = sc. parallelize (1,2,3,4,5,2,4,1)
	Comer = 40 purely (1) 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
	Teacher's Signature

POPU	
	Date
Topic	Paga No
2 cache the RDD. - content, cachec	,
-> context, cache()	
3 check is RDD is cooped on NOT	
3 check if RDD is cached or NOT.  -> content, is cached	
The same of the sa	
4 Remove Box dia Laura di visa di la	
y Remone persistance using impersion - content unpersist	FO
contest, conference	
- Check is ROD in contrad on him	
→ check if RDD is cached OR NDt	
Coruera o Cachia.	
6 Carlos Has DOO HAIDS boursely	
5 Cache the ROD using persist () -> content. persist ()	
- ontent perdiate	
7 classes is non in contract	
7 check if RDD is cached or not.	
-> content is coined.	
	1./ \
8 Remone persistance using unpersis - Content unpersist	PU
- Content . Unpersust,	
9 check if RDD is cached or Met.	
- content. es cached.	
10 cache the ROD when we have mul transformations.	tiple
transformations.	
Teacher's Signature .	

POPU°	
	Date
Topic	. Page No
-> a=sc. textfile ("dbys:/Filestore/table a. take (5)	es/mouils. +x+")
-> b = a.map (lambda x: x. upper ())	
b. take (5)	
-> c = b. filter (lambda a: x. Hartsu	rith ("1")
c. take(s)	
c. court.	
-> c. persiet	
-) C. Court	
Tambada Clamatu	ro.
Teacher's Signatus	Jo

-JAMA-

Owbus

§ (1) SpackTobs

6 (1) Spark John

autilis: [" ('space, 1)", "('sun', 1)", "('programs', 2)", "("in", 1)", "(clasy', 2)")

· Caching

2 OUT(15): Parallelcollection RDD [21) at readRDD From InputStream at Python ROD

3 True

4 OUT[12]: Parallecollection ROD[21] at read to O From Inputation at Bython ROD

False OUT[17]: Paraelillollection PDD (27) at read PDD From Inputstream at Mython ROD

7 True

& cound method ROD, unpersiat of Parallel Collection PO [21) at readrop from Input stream at Pythourid.

2 True

10 OUT[21]:
[(1: Toystory (1995): Animated | children's) comody's
"2: Jumanji (1995): Adventure | children's | Pantasy's

11 OUT [22]:

[(1: TOYSTORY (1995): ANIMATED | CHILDREN'S | COMEDY!

112: TUMANJI (1995): ADVENTURE) CHILDREN'S | FANTASY!

12 OUT[23]:
["1: TOYSTORY (1995): ANIMATED | CHILDREN'S | COMERY"
110: GOLDEN EYE (1995): ACTION) ADVENTURE | TURILLER"

THE REPORT OF THE PROPERTY OF

13 au (24]: 1055

14 Out [25]: Python 200 [28) at RDD at Python RDD

12 1022

when the restable.

and the security from file.

asser a such suing in walk signific