ASSIGNMENT 24.1

Task 1

Read a stream of Strings, fetch the words which can be converted to numbers. Filter out the rows, where the sum of numbers in that line is odd.
Provide the sum of all the remaining numbers in that batch.

SOLUTION-Task1

The source code of the program is given below. The program does the following

1. It creates a text file stream and monitors a directory for a text file where we supply it the words
2. Once the file is written, we iterate over all RDDs and its words
3. We created a function getDigit which returns 0 for non-digit chars and the digit for the number chars
4. For each string in the RDD we apply the same function which returns a vector of numbers and we use vector.sum to check if the sum is even or odd
5. For odd sum we ignore the numbers and for even sum we print the sum of the digits.
6. The program output is given after the source code

```
package Stream

import org.apache.spark.{SparkConf, SparkContext}
import org.apache.spark.streaming.{Seconds, StreamingContext}

object StreamingCaseStudy {
  def main(args: Array[String]): Unit = {


    if(args.length != 1)
    {
      println("Please provide local filename")
      return

    }
    val conf = new
SparkConf().setMaster("local[2]").setAppName("SparkSteamingExample")
    val sc = new SparkContext(conf)

    sc.setLogLevel("WARN")

    val ssc = new StreamingContext(sc, Seconds(5))
    println("Spark Streaming Context Created !")

    val lines = ssc.textFileStream("file://"+args(0))
    val words = lines.flatMap(_.split(" "))


    words.foreachRDD(a => {

      a.foreach(f => {

        if(f.map(x=>getDigits(x)).sum %2 == 0)
        {
          println("word is " + f + "  sum is " + f.map(x=>getDigits(x)).sum )
        }
        else
        {
```

```
          println("word is " + f + " ignored")
        }

      })

    });
    ssc.start()
    ssc.awaitTermination()


  }
```

The following strings are present the file from which the stream is read

```
[acadgild.mmisra words]$ cat words.txt
abcd9
aa112233
abcd11cdee22
ramu1579uget
filterthis113
passthis22222222
magic1356
[acadgild.mmisra words]$
```

Here is the output of the program corresponding to this stream of words

```
[acadgild.mmisra ~]$ spark-submit --class Stream.StreamingCaseStudy --deploy-mode
client \
> --master spark//$SPARK_MASTER_IP:$SPARK_MASTER_PORT /home/acadgild/spark1.jar
/home/acadgild/words
18/09/27 06:02:07 INFO spark.SparkContext: Running Spark version 2.2.1
18/09/27 06:02:08 WARN util.NativeCodeLoader: Unable to load native-hadoop library for
your platform... using builtin-java classes where applicable
18/09/27 06:02:08 WARN util.Utils: Your hostname, localhost.localdomain resolves to a
loopback address: 127.0.0.1; using 192.168.0.107 instead (on interface eth16)
18/09/27 06:02:08 WARN util.Utils: Set SPARK_LOCAL_IP if you need to bind to another
address
18/09/27 06:02:08 INFO spark.SparkContext: Submitted application: SparkSteamingExample
18/09/27 06:02:08 INFO spark.SecurityManager: Changing view acls to: acadgild
18/09/27 06:02:08 INFO spark.SecurityManager: Changing modify acls to: acadgild
18/09/27 06:02:08 INFO spark.SecurityManager: Changing view acls groups to:
18/09/27 06:02:08 INFO spark.SecurityManager: Changing modify acls groups to:
18/09/27 06:02:08 INFO spark.SecurityManager: SecurityManager: authentication
disabled; ui acls disabled; users  with view permissions: Set(acadgild); groups with
view permissions: Set(); users  with modify permissions: Set(acadgild); groups with
modify permissions: Set()
18/09/27 06:02:08 INFO util.Utils: Successfully started service 'sparkDriver' on port
42438.
18/09/27 06:02:08 INFO spark.SparkEnv: Registering MapOutputTracker
18/09/27 06:02:08 INFO spark.SparkEnv: Registering BlockManagerMaster
18/09/27 06:02:08 INFO storage.BlockManagerMasterEndpoint: Using
org.apache.spark.storage.DefaultTopologyMapper for getting topology information
18/09/27 06:02:08 INFO storage.BlockManagerMasterEndpoint: BlockManagerMasterEndpoint
up
18/09/27 06:02:09 INFO storage.DiskBlockManager: Created local directory at
/tmp/blockmgr-82a1437f-cd8b-4a59-9ba2-32278997ee71
18/09/27 06:02:09 INFO memory.MemoryStore: MemoryStore started with capacity 366.3 MB
18/09/27 06:02:09 INFO spark.SparkEnv: Registering OutputCommitCoordinator
18/09/27 06:02:09 INFO util.log: Logging initialized @2439ms
18/09/27 06:02:09 INFO server.Server: jetty-9.3.z-SNAPSHOT
18/09/27 06:02:09 INFO server.Server: Started @2579ms
18/09/27 06:02:09 INFO server.AbstractConnector: Started
```

```
ServerConnector@3abfc4ed{HTTP/1.1,[http/1.1]}{0.0.0.0:4040}
18/09/27 06:02:09 INFO util.Utils: Successfully started service 'SparkUI' on port
4040.
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@70e29e14{/jobs,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@1a15b789{/jobs/json,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@51650883{/jobs/job,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@30c31dd7{/jobs/job/json,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@596df867{/stages,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@241a53ef{/stages/json,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@2db2cd5{/stages/stage,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@73393584{/stages/stage/json,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@1827a871{/stages/pool,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@7249dadf{/stages/pool/json,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@66238be2{/storage,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@200606de{/storage/json,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@f8908f6{/storage/rdd,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@2ef8a8c3{/storage/rdd/json,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@63fd4873{/environment,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@7544a1e4{/environment/json,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@7957dc72{/executors,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@3aacf32a{/executors/json,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@82c57b3{/executors/threadDump,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@600b0b7{/executors/threadDump/json,null,AVAILABLE,@Spark
}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@5ea502e0{/static,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@d816dde{/,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@6c451c9c{/api,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@4e9658b5{/jobs/job/kill,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@20312893{/stages/stage/kill,null,AVAILABLE,@Spark}
18/09/27 06:02:09 INFO ui.SparkUI: Bound SparkUI to 0.0.0.0, and started at
http://192.168.0.107:4040
18/09/27 06:02:09 INFO spark.SparkContext: Added JAR file:/home/acadgild/spark1.jar at
spark://192.168.0.107:42438/jars/spark1.jar with timestamp 1538008329579
18/09/27 06:02:09 INFO executor.Executor: Starting executor ID driver on host
localhost
18/09/27 06:02:09 INFO util.Utils: Successfully started service
'org.apache.spark.network.netty.NettyBlockTransferService' on port 46875.
18/09/27 06:02:09 INFO netty.NettyBlockTransferService: Server created on
```

```
192.168.0.107:46875
18/09/27 06:02:09 INFO storage.BlockManager: Using
org.apache.spark.storage.RandomBlockReplicationPolicy for block replication policy
18/09/27 06:02:09 INFO storage.BlockManagerMaster: Registering BlockManager
BlockManagerId(driver, 192.168.0.107, 46875, None)
18/09/27 06:02:09 INFO storage.BlockManagerMasterEndpoint: Registering block manager
192.168.0.107:46875 with 366.3 MB RAM, BlockManagerId(driver, 192.168.0.107, 46875,
None)
18/09/27 06:02:09 INFO storage.BlockManagerMaster: Registered BlockManager
BlockManagerId(driver, 192.168.0.107, 46875, None)
18/09/27 06:02:09 INFO storage.BlockManager: Initialized BlockManager:
BlockManagerId(driver, 192.168.0.107, 46875, None)
18/09/27 06:02:10 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@1813f3e9{/metrics/json,null,AVAILABLE,@Spark}
Spark Streaming Context Created !
word is abcd9 ignored
word is aa112233  sum is 12
word is abcd11cdee22  sum is 6
word is ramu1579uget  sum is 22
word is filterthis113 ignored
word is passthis22222222  sum is 16
word is magic1356 ignored
```

Task 2

Read two streams

1. List of strings input by user
2. Real-time set of offensive words

Find the word count of the offensive words inputted by the user as per the real-time set of offensive words

Solution-Task2

The source code for task2 is the given below. The program does the following

1. Open one stream to read list of bad words from a file in a local directory
2. Opens another stream to receive user input from socket at port 9000
3. When we place a file with bad words in the directory, the program reads the stream and stores each bad word in a ListBuffer
4. When user input is received from other stream, we first do the word count of the words input by the user. Then we filter the RDD for all the words matching with any of the words in the bad word list.
5. Once filtered, we show the count of bad words to the user along with the input

```scala
package Stream


import org.apache.spark.{SparkConf, SparkContext}
import org.apache.spark.streaming.{Seconds, StreamingContext}

import scala.collection.mutable
import scala.collection.mutable.{ArrayBuffer, ListBuffer}

object SparkTwoStream {

  val abuses:ListBuffer[String] = ListBuffer.empty[String]
  def main(args: Array[String]): Unit = {


    if(args.length != 1)
      {
        println("provide local directory for reading offensive words")
      }
    // create a list of abuses that we keep storing in memory based on what we receive
from the offensive words stream


    val conf = new
SparkConf().setMaster("local[2]").setAppName("SparkSteamingExample")
    val sc = new SparkContext(conf)

    sc.setLogLevel("WARN")

    val ssc = new StreamingContext(sc, Seconds(5))
    println("Spark Streaming Context Created !")

    // create stream for offensive words to be read from a file
    val lines = ssc.textFileStream("file://"+args(0))
    val words = lines.flatMap(_.split(" "))

    words.foreachRDD(a => {

      a.foreach(f => {

        // prepend the list with the offensive word

        abuses+= f.toLowerCase()

      })

    })


    val input = ssc.socketTextStream("localhost", 9000)
    //Getting word count of all words entered by user
    val userwords = input.flatMap(line => line.split(" ")).map(word => (word,
1)).reduceByKey(_ + _);

    val userbadwords = userwords.filter{x=>abuses.contains(x._1.toLowerCase())}

    // print user input
    input.print()
    // print count of bad words in input
    userbadwords.print()
```

```
    ssc.start()
    ssc.awaitTermination()


  }


}
```

Program Output

Steps

We start the netcat server on port 9000 for the user input

```
You have mail in /var/spool/mail/acadgild
[acadgild.mmisra ~]$ nc -lk 9000
I am very small
I am not very good
I am an stupid person
You are idiot stupid and chutia person
I am idiot idiot but not stupid stupid
I am pagal but not ullu
```

Launch Spark application with spark submit utility. We also provide the local directory (/home/acadgild/words) to monitor to the program

```
Last login: Fri Sep 28 15:42:12 2018 from 192.168.56.1
[acadgild.mmisra ~]$ spark-submit --class Stream.SparkTwoStream --deploy-mode client \
> --master spark//$SPARK_MASTER_IP:$SPARK_MASTER_PORT /home/acadgild/spark1.jar
/home/acadgild/words
18/09/28 15:46:03 INFO spark.SparkContext: Running Spark version 2.2.1
18/09/28 15:46:03 WARN util.NativeCodeLoader: Unable to load native-hadoop library for
your platform... using builtin-java classes where applicable
18/09/28 15:46:03 WARN util.Utils: Your hostname, localhost.localdomain resolves to a
loopback address: 127.0.0.1; using 10.0.2.15 instead (on interface eth17)
18/09/28 15:46:03 WARN util.Utils: Set SPARK_LOCAL_IP if you need to bind to another
address
18/09/28 15:46:03 INFO spark.SparkContext: Submitted application: SparkSteamingExample
18/09/28 15:46:03 INFO spark.SecurityManager: Changing view acls to: acadgild
18/09/28 15:46:03 INFO spark.SecurityManager: Changing modify acls to: acadgild
18/09/28 15:46:03 INFO spark.SecurityManager: Changing view acls groups to:
18/09/28 15:46:03 INFO spark.SecurityManager: Changing modify acls groups to:
18/09/28 15:46:03 INFO spark.SecurityManager: SecurityManager: authentication
disabled; ui acls disabled; users  with view permissions: Set(acadgild); groups with
view permissions: Set(); users  with modify permissions: Set(acadgild); groups with
modify permissions: Set()
18/09/28 15:46:04 INFO util.Utils: Successfully started service 'sparkDriver' on port
41732.
18/09/28 15:46:04 INFO spark.SparkEnv: Registering MapOutputTracker
18/09/28 15:46:04 INFO spark.SparkEnv: Registering BlockManagerMaster
18/09/28 15:46:04 INFO storage.BlockManagerMasterEndpoint: Using
org.apache.spark.storage.DefaultTopologyMapper for getting topology information
18/09/28 15:46:04 INFO storage.BlockManagerMasterEndpoint: BlockManagerMasterEndpoint
up
18/09/28 15:46:04 INFO storage.DiskBlockManager: Created local directory at
/tmp/blockmgr-181e7e3c-a3c2-44d9-b1b1-85a439df8ce5
```

```
18/09/28 15:46:04 INFO memory.MemoryStore: MemoryStore started with capacity 366.3 MB
18/09/28 15:46:04 INFO spark.SparkEnv: Registering OutputCommitCoordinator
18/09/28 15:46:04 INFO util.log: Logging initialized @2480ms
18/09/28 15:46:04 INFO server.Server: jetty-9.3.z-SNAPSHOT
18/09/28 15:46:04 INFO server.Server: Started @2595ms
18/09/28 15:46:04 INFO server.AbstractConnector: Started
ServerConnector@63a564f2{HTTP/1.1,[http/1.1]}{0.0.0.0:4040}
18/09/28 15:46:04 INFO util.Utils: Successfully started service 'SparkUI' on port
4040.
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@7646731d{/jobs,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@797501a{/jobs/json,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@57f791c6{/jobs/job,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@5bd1ceca{/jobs/job/json,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@499b2a5c{/stages,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@c1fca1e{/stages/json,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@344344fa{/stages/stage,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@285f09de{/stages/stage/json,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@31500940{/stages/pool,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@48e64352{/stages/pool/json,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@4362d7df{/storage,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@1c25b8a7{/storage/json,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@750fe12e{/storage/rdd,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@3e587920{/storage/rdd/json,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@24f43aa3{/environment,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@1e11bc55{/environment/json,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@70e0accd{/executors,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@6ab72419{/executors/json,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@4fdfa676{/executors/threadDump,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@5be82d43{/executors/threadDump/json,null,AVAILABLE,@Spar
k}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@345e5a17{/static,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@1e886a5b{/,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@6e33c391{/api,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@213e3629{/jobs/job/kill,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@2a7b6f69{/stages/stage/kill,null,AVAILABLE,@Spark}
18/09/28 15:46:04 INFO ui.SparkUI: Bound SparkUI to 0.0.0.0, and started at
http://10.0.2.15:4040
18/09/28 15:46:04 INFO spark.SparkContext: Added JAR file:/home/acadgild/spark1.jar at
```

```
spark://10.0.2.15:41732/jars/spark1.jar with timestamp 1538129764779
18/09/28 15:46:04 INFO executor.Executor: Starting executor ID driver on host
localhost
18/09/28 15:46:04 INFO util.Utils: Successfully started service
'org.apache.spark.network.netty.NettyBlockTransferService' on port 35949.
18/09/28 15:46:04 INFO netty.NettyBlockTransferService: Server created on
10.0.2.15:35949
18/09/28 15:46:04 INFO storage.BlockManager: Using
org.apache.spark.storage.RandomBlockReplicationPolicy for block replication policy
18/09/28 15:46:04 INFO storage.BlockManagerMaster: Registering BlockManager
BlockManagerId(driver, 10.0.2.15, 35949, None)
18/09/28 15:46:05 INFO storage.BlockManagerMasterEndpoint: Registering block manager
10.0.2.15:35949 with 366.3 MB RAM, BlockManagerId(driver, 10.0.2.15, 35949, None)
18/09/28 15:46:05 INFO storage.BlockManagerMaster: Registered BlockManager
BlockManagerId(driver, 10.0.2.15, 35949, None)
18/09/28 15:46:05 INFO storage.BlockManager: Initialized BlockManager:
BlockManagerId(driver, 10.0.2.15, 35949, None)
18/09/28 15:46:05 INFO handler.ContextHandler: Started
o.s.j.s.ServletContextHandler@375b5b7f{/metrics/json,null,AVAILABLE,@Spark}
Spark Streaming Context Created !
-------------------------------------------
Time: 1538129770000 ms
-------------------------------------------

-------------------------------------------
```

We create a file called badwords.txt in the /home/acadgild/words directory with the following bad words

```
[acadgild.mmisra words]$ vim badwords.txt
You have mail in /var/spool/mail/acadgild
[acadgild.mmisra words]$ ls -la
total 24
drwxrwxr-x.  2 acadgild acadgild  4096 Sep 28 15:48 .
drwx------. 71 acadgild acadgild 12288 Sep 28 15:48 ..
-rw-rw-r--.  1 acadgild acadgild    26 Sep 28 15:47 badwords.txt
 [acadgild.mmisra words]$ cat badwords.txt
stupid
idiot
moron
chutia
```

We provide user input on the netcat window with the following sentences and see the program output

```
[acadgild.mmisra ~]$ nc -lk 9000
I am very small
I am not very good
I am an stupid person
You are idiot stupid and chutia person
I am idiot idiot but not stupid stupid

-------------------------------------------
Time: 1538129785000 ms
-------------------------------------------

18/09/28 15:46:26 WARN storage.RandomBlockReplicationPolicy: Expecting 1 replicas with
only 0 peer/s.
18/09/28 15:46:26 WARN storage.BlockManager: Block input-1-1538129786600 replicated to
```

```
only 0 peer(s) instead of 1 peers
-------------------------------------------
Time: 1538129790000 ms
-------------------------------------------
I am not very good

-------------------------------------------
Time: 1538129790000 ms
-------------------------------------------

18/09/28 15:46:31 WARN storage.RandomBlockReplicationPolicy: Expecting 1 replicas with
only 0 peer/s.
18/09/28 15:46:31 WARN storage.BlockManager: Block input-1-1538129790800 replicated to
only 0 peer(s) instead of 1 peers
-------------------------------------------
Time: 1538129795000 ms
-------------------------------------------
I am an stupid person

-------------------------------------------
Time: 1538129795000 ms
-------------------------------------------
(stupid,1)

-------------------------------------------
Time: 1538129800000 ms
-------------------------------------------


18/09/28 15:46:56 WARN storage.RandomBlockReplicationPolicy: Expecting 1 replicas with
only 0 peer/s.
18/09/28 15:46:56 WARN storage.BlockManager: Block input-1-1538129815800 replicated to
only 0 peer(s) instead of 1 peers
-------------------------------------------
Time: 1538129820000 ms
-------------------------------------------
You are idiot stupid and chutia person

-------------------------------------------
Time: 1538129820000 ms
-------------------------------------------
(chutia,1)
(stupid,1)
(idiot,1)

-------------------------------------------
Time: 1538129825000 ms
-------------------------------------------

18/09/28 15:47:23 WARN storage.RandomBlockReplicationPolicy: Expecting 1 replicas with
only 0 peer/s.
18/09/28 15:47:23 WARN storage.BlockManager: Block input-1-1538129843000 replicated to
only 0 peer(s) instead of 1 peers
-------------------------------------------
Time: 1538129845000 ms
-------------------------------------------
I am idiot idiot but not stupid stupid

-------------------------------------------
Time: 1538129845000 ms
-------------------------------------------
(stupid,2)
(idiot,2)
```

```
-------------------------------------------
Time: 1538129850000 ms
-------------------------------------------
```

We create one more file called morebadwords.txt in the /home/acadgild/words directory with the following bad words. This will add two more bad words in the ListBuffer (dictionary)

```
[acadgild.mmisra words]$ vim morebadwords.txt
[acadgild.mmisra words]$ ls -la
total 24
drwxrwxr-x.  2 acadgild acadgild  4096 Sep 28 15:48 .
drwx------. 71 acadgild acadgild 12288 Sep 28 15:48 ..
-rw-rw-r--.  1 acadgild acadgild    26 Sep 28 15:47 badwords.txt
-rw-rw-r--.  1 acadgild acadgild    11 Sep 28 15:48 morebadwords.txt
[acadgild.mmisra words]$ cat morebadwords.txt
pagal
ullu
[acadgild.mmisra words]$
```

We type one more user input see that input text is validated against the newly added bad words

```
[acadgild.mmisra ~]$ nc -lk 9000
I am very small
I am not very good
I am an stupid person
You are idiot stupid and chutia person
I am idiot idiot but not stupid stupid
I am pagal but not ullu


-------------------------------------------
Time: 1538129915000 ms
-------------------------------------------
I am pagal but not ullu

-------------------------------------------
Time: 1538129915000 ms
-------------------------------------------
(ullu,1)
(pagal,1)

-------------------------------------------
Time: 1538129920000 ms
-------------------------------------------

-------------------------------------------
Time: 1538129920000 ms
-------------------------------------------
```