WORDCOUNT

package wordcount;

import org.apache.hadoop.conf.Configured;

import org.apache.hadoop.fs.Path;

import org.apache.hadoop.io.IntWritable;

import org.apache.hadoop.io.Text;

import org.apache.hadoop.mapred.FileInputFormat;

import org.apache.hadoop.mapred.FileOutputFormat;

import org.apache.hadoop.mapred.JobClient;

import org.apache.hadoop.mapred.JobConf;

import org.apache.hadoop.util.Tool;

import org.apache.hadoop.util.ToolRunner;

public class WordCount extends Configured implements Tool{

public int run(String[] args)throws Exception{

if(args.length<2)

{

System.out.println("Please add proper input and output directory");

return -1;

}

JobConf conf=new JobConf(WordCount.class);

FileInputFormat.setInputPaths(conf,new Path(args[0]));

FileOutputFormat.setOutputPath(conf,new Path(args[1]));

conf.setMapperClass(WordMapper.class);

conf.setReducerClass(WordReducer.class);

conf.setOutputKeyClass(Text.class);

conf.setMapOutputValueClass(IntWritable.class);

conf.setOutputKeyClass(Text.class);

conf.setOutputValueClass(IntWritable.class);

JobClient.runJob(conf);

return 0;

}

public static void main(String[] args) throws Exception{

int exitCode = ToolRunner.run(new WordCount(),args);

System.exit(exitCode);

}

}

WORDMAPPER

package wordcount;

import java.io.IOException;

import org.apache.hadoop.io.IntWritable;

import org.apache.hadoop.io.LongWritable;

import org.apache.hadoop.io.Text;

import org.apache.hadoop.mapred.MapReduceBase;

import org.apache.hadoop.mapred.OutputCollector;

import org.apache.hadoop.mapred.Reporter;

import org.apache.hadoop.mapred.Mapper;

public class WordMapper extends MapReduceBase implements Mapper<LongWritable, Text,Text,IntWritable>

{

public void map(LongWritable key,Text value, OutputCollector<Text,IntWritable>output,Reporter r) throws IOException{

String s=value.toString();

for(String word:s.split(" "))

{

if(word.length()>0)

{

output.collect(new Text(word), new IntWritable(1));

}

}

}

}

WORDREDUCER

package wordcount;

import java.io.IOException;

import org.apache.hadoop.io.IntWritable;

import org.apache.hadoop.io.LongWritable;

import org.apache.hadoop.io.Text;

import org.apache.hadoop.mapred.MapReduceBase;

import org.apache.hadoop.mapred.OutputCollector;

import org.apache.hadoop.mapred.Reporter;

import org.apache.hadoop.mapred.Mapper;

public class WordMapper extends MapReduceBase implements Mapper<LongWritable, Text,Text,IntWritable>

{

public void map(LongWritable key,Text value, OutputCollector<Text,IntWritable>output,Reporter r) throws IOException{

String s=value.toString();

for(String word:s.split(" "))

{

if(word.length()>0)

{

output.collect(new Text(word), new IntWritable(1));

}

}

}

}