Big Data Programming Extra Credit 1

Shayideep Sangam

Code:

```
package spark10;
import java.util.Arrays;
import java.util.List;

import org.apache.spark.api.java.JavaRDD;
import org.apache.spark.api.java.JavaSparkContext;
import org.apache.spark.mllib.fpm.AssociationRules;
import org.apache.spark.mllib.fpm.FPGrowth;
import org.apache.spark.mllib.fpm.FPGrowthModel;
import org.apache.spark.sql.SparkSession;
public class frequentpair {
	public static void main(String[] args) {
	SparkSession spark =
SparkSession.builder().config("spark.master","local[*]").getOrCreate();
	JavaSparkContext sc = new JavaSparkContext(spark.sparkContext());
	//sc.setLogLevel("WARN");
	JavaRDD<String> data =
sc.textFile("file:///C:/Users/dell/Downloads/sample_fpgrowth.txt");

	JavaRDD<List<String>> transactions = data.map(line ->
Arrays.asList(line.split(" ")));

	FPGrowth fpg = new FPGrowth()
	  .setMinSupport(0.2)
	  .setNumPartitions(10);

	FPGrowthModel<String> model = fpg.run(transactions);

		for (FPGrowth.FreqItemset<String> itemset :
model.freqItemsets().toJavaRDD().collect()) {
			System.out.println("[" + itemset.javaItems() + "], " +
itemset.freq());
		}

	double minConfidence = 0.8;
			for (AssociationRules.Rule<String> rule
			  :
model.generateAssociationRules(minConfidence).toJavaRDD().collect()) {
			System.out.println(
			  rule.javaAntecedent() + " => " + rule.javaConsequent()
+ ", " + rule.confidence());
			}
}
}
```

Output:

Frequent Patterns:

```
[[17]], 5
[[15]], 4
[[15, 17]], 3
[[16]], 3
[[16, 15]], 3
[[16, 15, 17]], 3
[[16, 17]], 3
[[9]], 3
[[9, 15]], 2
[[9, 17]], 2
[[10]], 3
[[10, 16]], 2
[[10, 16, 15]], 2
[[10, 16, 15, 17]], 2
[[10, 16, 17]], 2
[[10, 15]], 3
[[10, 15, 17]], 2
[[10, 17]], 2
[[11]], 3
[[11, 16]], 3
[[11, 16, 15]], 3
[[11, 16, 15, 17]], 3
[[11, 16, 17]], 3
[[11, 10]], 2
[[11, 10, 16]], 2
[[11, 10, 16, 15]], 2
[[11, 10, 16, 15, 17]], 2
[[11, 10, 16, 17]], 2
[[11, 10, 15]], 2
[[11, 10, 15, 17]], 2
[[11, 10, 17]], 2
[[11, 15]], 3
[[11, 15, 17]], 3
[[11, 17]], 3
[[7]], 2
[[11, 17]], 3
[[7]], 2
[[7, 9]], 2
[[7, 9, 17]], 2
[[7, 17]], 2
[[8]], 2
[[8, 16]], 2
[[8, 16, 15]], 2
[[8, 16, 15, 17]], 2
[[8, 16, 17]], 2
[[8, 11]], 2
[[8, 11, 16]], 2
[[8, 11, 16, 15]], 2
[[8, 11, 16, 15, 17]], 2
[[8, 11, 16, 17]], 2
[[8, 11, 15]], 2
[[8, 11, 15, 17]], 2
[[8, 11, 17]], 2
[[8, 15]], 2
[[8, 15, 17]], 2
[[8, 17]], 2
18/11/24 16:01:16 INFO Spa
18/11/24 16:01:16 INFO Mar
```

Association Rules:

```
[11, 16, 17] => [15], 1.0
[8, 11, 15, 17] => [16], 1.0
[8, 11, 16, 17] => [15], 1.0
[10, 16, 17] => [15], 1.0
[10, 16, 17] => [11], 1.0
[8] => [16], 1.0
[8] => [11], 1.0
[8] => [15], 1.0
[8] => [17], 1.0
[8, 11, 17] => [16], 1.0
[8, 11, 17] => [15], 1.0
[16, 17] => [15], 1.0
[16, 17] => [11], 1.0
[8, 16, 15, 17] => [11], 1.0
[11, 10, 15, 17] => [16], 1.0
[10, 15, 17] => [16], 1.0
[10, 15, 17] => [11], 1.0
[8, 15] => [16], 1.0
[8, 15] => [11], 1.0
[8, 15] => [17], 1.0
[11, 10, 16, 17] => [15], 1.0
[11, 10, 16, 15] => [17], 1.0
[11, 16, 15] => [17], 1.0
[16, 15] => [17], 1.0
[16, 15] => [11], 1.0
[8, 15, 17] => [16], 1.0
[8, 15, 17] => [11], 1.0
[10, 16] => [15], 1.0
[10, 16] => [17], 1.0
[10, 16] => [11], 1.0
[11, 15, 17] => [16], 1.0
[8, 11] => [16], 1.0
[8, 11] => [15], 1.0
[8, 11] => [17], 1.0
[8, 11, 16] => [15], 1.0
[8, 11, 16] => [17], 1.0
[15, 17] => [16], 1.0
[15, 17] => [11], 1.0
[7] => [9], 1.0
[7] => [17], 1.0
[8, 16] => [15], 1.0
[8, 16] => [17], 1.0
[8, 16] => [11], 1.0
[7, 17] => [9], 1.0
[8, 16, 17] => [15], 1.0
[8, 16, 17] => [11], 1.0
[8, 11, 16, 15] => [17], 1.0
[16] => [15], 1.0
[16] => [17], 1.0
[16] => [11], 1.0
[10, 17] => [16], 1.0
[10, 17] => [15], 1.0
[10, 17] => [11], 1.0
[10, 16, 15] => [17], 1.0
[10, 16, 15] => [11], 1.0
[8, 17] => [16], 1.0
[8, 17] => [11], 1.0
[8, 17] => [15], 1.0
[8, 16, 15] => [17], 1.0
[8, 16, 15] => [11], 1.0
[16, 15, 17] => [11], 1.0
[10, 16, 15, 17] => [11], 1.0
[11, 15] => [16], 1.0
[11, 15] => [17], 1.0
[11] => [16], 1.0
[11] => [15], 1.0
[11] => [17], 1.0
[11, 10, 16] => [15], 1.0
[11, 10, 16] => [17], 1.0
[11, 16] => [15], 1.0
[11, 16] => [17], 1.0
[11, 10] => [16], 1.0
11, 10] => [16], 1.0
11, 10] => [15], 1.0
11, 10] => [17], 1.0
7, 9] => [17], 1.0
10] => [15], 1.0
11, 17] => [16], 1.0
11, 17] => [15], 1.0
9, 17] => [7], 1.0
8, 11, 15] => [16], 1.0
8, 11, 15] => [17], 1.0
11, 10, 15] => [16], 1.0
11, 10, 15] => [17], 1.0
11, 10, 17] => [16], 1.0
11, 10, 17] => [15], 1.0
8/11/24 16:01:20 INFO SparkC
8/11/24 16:01:20 INFO SparkU
```