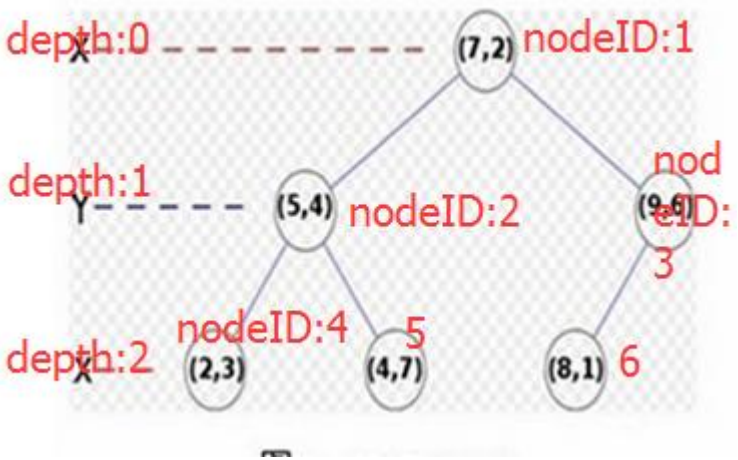
Kdtree spark design

If you have any question or suggestions,please contact me,thank you!

(1016830193@qq.com)

- Build example: (2 dim.)
- Input data: {（2,3），（5,4），（9,6），（4,7），（8,1），（7,2）}



- ps(x:dim.0, Y:dim.1)
- Parallel ideas:
- From depth 0->1: select (7,2) ,dim.0 to split data as (2,((5,4),(2,3),(4,7))), (3,((9,6),(8,1))) ,then groupByKey(),flatMap ,2 parts distribute process ,each part just as depth 0 ways
- So, after that, the result is: ((1,(7,2)), (2,(5,4)), (3,(9,6)), (4,(2,3)), (5,(4,7)), (6,(8,1)))
- …..

- key codes
- The overall cycle : update rootNode

```
36  while (currentParentNodes.last.id < dataLength){
37    keys = createKDTree(labeledInput, keys, depth , this.dim, dataLength, rootNode, currentParentNodes)
38    depth = depth + 1
39    updateCurrentParentNodes(currentParentNodes)
40  }
```

- This is in def createKDTree, can update nodeID,

```
75      val input = labeledInput.sparkContext.parallelize(keys).zip(labeledInput)
76      val keyPoints  = input.groupByKey()
77        .flatMap(nodePoints =>
78        updateKeys(nodePoints, currentParentNodes)
79        ).sortByKey()
80      val newKeyPoints = keyPoints.values.collect()
```