BigData Assignment 3.5

Write a hive UDF that implements functionality of string concat\_ws(string SEP, array<string>).This UDF will accept two arguments, one string and one array of string.It will return a single string where all the elements of the array are separated by the SEP.

Solution -

**StringUDF.java**

**package hive;**

**import java.util.ArrayList;**

**import org.apache.hadoop.hive.ql.exec.UDF;**

**public class StringUDF extends UDF{**

**public String evaluate(String sep,ArrayList<String> arr) {**

**String res = "";**

**if(sep==null || arr==null) {**

**//To check input is null or not**

**return null;**

**}**

**for(int i=0;i<arr.size()-1;i++) {**

**res = res+arr.get(i)+sep;**

**//To add separator**

**}**

**res = res+arr.get(arr.size()-1);**

**return res;**

**}**

**}**

* First started the all the required daemons

**start-all.sh**

* Started the hive shell

**hive**

* Checked the database and the content of the table

**use default;**

**select \* from student;**



* Added the hive jar for string separation

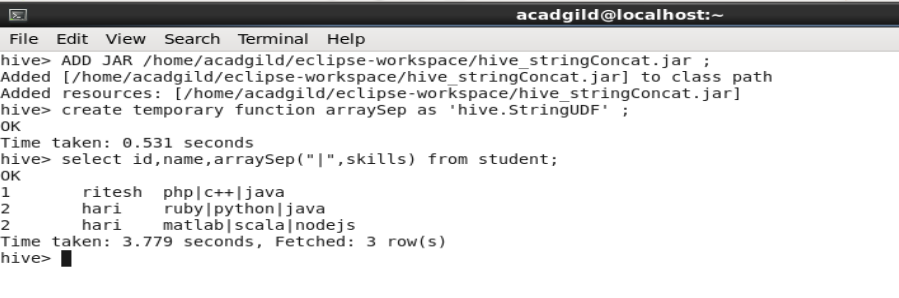
**ADD jar /home/acadgild/eclipse/hive\_stringConcat.jar;**

* Created a hive function t o use hive custom UDF

**create temporary function arraySep as 'hive.StringUDF';**

* Separated the skills array from the table using separator '|' and using hive custom UDF

**select id,name,arraySep(“|”,skills) from student;**



Output -

In the above screenshot , it is evident that we using hive custom UDF function , we added separator '|' in the skills array using that function.