**HƯỚNG DẪN THỰC THI DEMO**

Download máy ảo trong VirtualMachine từ link:

Import 1 master và 2 slave vào VMWare.

1. **Thực thi với Apriori trên môi trường song song**

Bước 1: Khởi động Hadoop bằng lệnh:

*start-dfs.sh*

*start-yarn.sh*

Bước 2: Đưa data lên hdfs:

*hdfs dfs -put dataSet/sample*

Bước 3: Thực thi với tập data *sample* bằng lệnh sau*:*

*hadoop jar HadoopBasedApriori.jar apriori.AprioriDriver sample output 3 0.5 6*

Trong đó:

*sample* là tập data trên hdfs

*3* là số itemset

*output* là thư mục lưu các tập output khi được sinh ra

*0.5* là độ support

*6* là số transaction

Bước 4: Xem output chương trình bằng lệnh:

*hdfs dfs –cat/home/khanhduyuser/output1/part-r-00000*

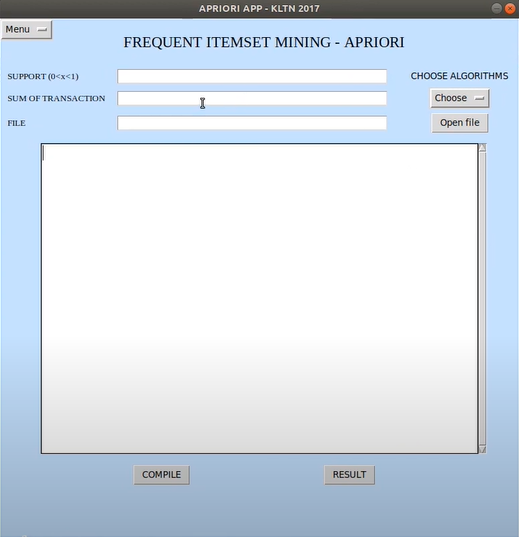
1. **Thực thi Apriori trên ứng dụng**

Bước 1: Khởi động Hadoop bằng lệnh:

*start-dfs.sh và start-yarn.sh*

Bước 2: Khới động app bằng lệnh :

*python3 apriori.py*



Hình 1. Giao diện sau khi khởi động

1. Support là chứa độ support ta cần nhập.
2. Sum of transaction: tổng số giao dịch của tập data.
3. File: đường dẫn đến file data cần thực thi. Chọn button Open File để lấy đường dẫn đến file data.
4. Button choose : lựa chọn thuật toán muốn thực thi.

Bao gồm :

+ Apriori tuần tự

+ Apriori tuần tự cải tiến

+ Apriori song song

1. Button compile: Compile thuật toán sau khi chọn các thông số ở trên .
2. Button result: xem kết quả sau khi thực thi.

Bước 3: Nhập các thông số cần thiết support, transaction, file data, chọn thuật toán.



Bước 4: Compile sau thi chọn ở trên.

Bước 5: Xem kết qua bằng button result.

