**Pseudocode**

Hadoop Mapper and Reducer for Stripes-Approach

2.a mapper\_stripes.py

---------------------------------

For each line in sys.stdin:

Split the line on ‘\t’ to get the userID, movieID and rating

If rating is greater than 4

Insert all the movieID’s corresponding to that userID in a dictionary

For each userID in the dictionary

Create a movie list, sorted by the movieID’s

for i in range ( 0 , len ( sorted\_movielist ) ) :

for j in range ( i + 1, len ( sorted\_movielist ) ) :

if the movieID’s are not same

if movieID1 is less than movieID2

make pairs of movieID1 and movieID2

else

make pairs of movieID2 and movieID1

for every pair occurring in the movie dictionary

increase the count for the movie occurrence

for every pair in movie dictionary

make movie stripe with the associative array (dictionary inside dictionary) along with the counts of the movies occurrences

for every pair in the movieStripe dictionary

emit the movieStripe as key-value pairs

2.b reducer\_stripes.py

---------------------------------

Read each line of mapper output as input

Split the line on ‘\t’

Create a movieDictionary and insert the corresponding movies

For each stripe in movieDictionary having the counts

Sum the counts

Element wise sum

For each pair in dictionary

Split the pair on ‘,’

If pair in dictionary has the count > 200

emit movieID1, movieID2, count

3.For matching the movieID’s with the movie names

Movie\_names.py

-----------------------------

Open the file u.item by importing the file in python program

Give the movieID’s along with the count as an input

For each line in movieFile u.item

Split the line on ‘|’ to get movieID and movie\_name

Store the movieID and movie\_name as key-value pairs

If rating > 200

Emit movie\_names for the corresponding movieID’s along with their counts as values