

به نام خدا

محمد عذیری

220797072

طبق تعریفی که از MR یا نگاشت کاهش داریم می‌دانیم که در مرحله اول باید مرحله map کردن را انجام بدهیم.

در این سوال ما باید طبق یک معیار فاصله، فاصله محصول مورد نظر را با تمامی محصولات بدست آوریم. برای این مثال معیار Jaccard Distance می‌تواند گزینه مناسبی باشد.

ما تمامی محصولات را به صورتی که مقدار key ما برابر خود محصول بوده و مقدار value ما برابر JD از محصول مورد نظر بوده مپ می‌کنیم.

می‌دانیم که میشود به صورت Sort شده این مقدار را نگه داشت. پس مشکلی در قسمت Reduce نیز نداریم و می‌توانیم k محصول را که کمترین JD را دارند انتخاب کنیم و با این روش ما k نزدیک ترین سبد به سبد مورد نظرمان را بدست آوریم.

در این سوال میتوانیم قسمت map را به صورت جداگانه روی تمامی سرور ها پیاده کنیم و در ادامه اگر توانایی ذخیره مقادیر map شده در یک جا وجود داشت به صورت بالا عمل کنیم و در غیر این صورت برای هر سرور کل عملیات بالا را انجام دهیم و در اخر از بین 5k محصول بدست آمده k نزدیک ترین هارا انتخاب کنیم.