

## به نام خدا

### تمرین آشنایی با هادوپ و برنامه نویسی مپ ریدیوس دانشگاه صنعتی اصفهان - ترم اول ۱۴۰۱

در فایل پاسخ تکلیف، روند اجرای هر بخش و نتیجه را گزارش دهید. همچنین در ابتدای فایل، مشخص کنید از چه بستری برای انجام تکلیف استفاده کرده اید. همچنین از یک بار اجرای کامل تکلیف، رکوردی از دسکتاپ گرفته و همراه با فایل متنی پاسخ تکلیف، در یکتا آپلود کنید

(۱) یک برنامه ساده word count نوشته (جاوا یا پایتون)

(ب) برنامه را با استفاده از سیستم خود برای یک فایل ورودی حدود ۶۴ مگابایت اجرا کنید (این فایل را می‌توانید از روی اینترنت دانلود کنید و یا با ایجاد دیکشنری از کلمات محدود و تکرار آنها به صورت تصادفی در یک فایل جدید بسازید). به لاگ خروجی اجرا دقت کنید. چند map و چند reduce اجرا می‌شود؟

(ج) از فایل ورودی ۱۰ کپی بگیرید و برنامه را برای ۱۰ فایل به صورت همزمان اجرا کنید. چند map و چند reduce اجرا می‌شود؟  
(د) ۱۰ فایل را در یک فایل کپی کنید تا یک فایل حدود ۶۴۰ مگابایتی به دست آید. برنامه را برای این فایل جدید اجرا کنید. چند map و چند reduce اجرا می‌شود؟

(ه) به روش قبل، فایلی در حدود ۳۰۰ مگابایت بسازید و بررسی کنید چند map و چند reduce برای اجرای آن ساخته می‌شود.  
(و) سعی کنید با تنظیمات یا با اضافه کردن دستورات لازم به کد درایور، تعداد map و reduce ها را برای فایل ۳۰۰ مگابایتی تغییر دهید.

(ز) تحقیق کنید در حالت پیش فرض تعداد map های ساخته شده چه ارتباطی با اندازه فایل ورودی دارد.