

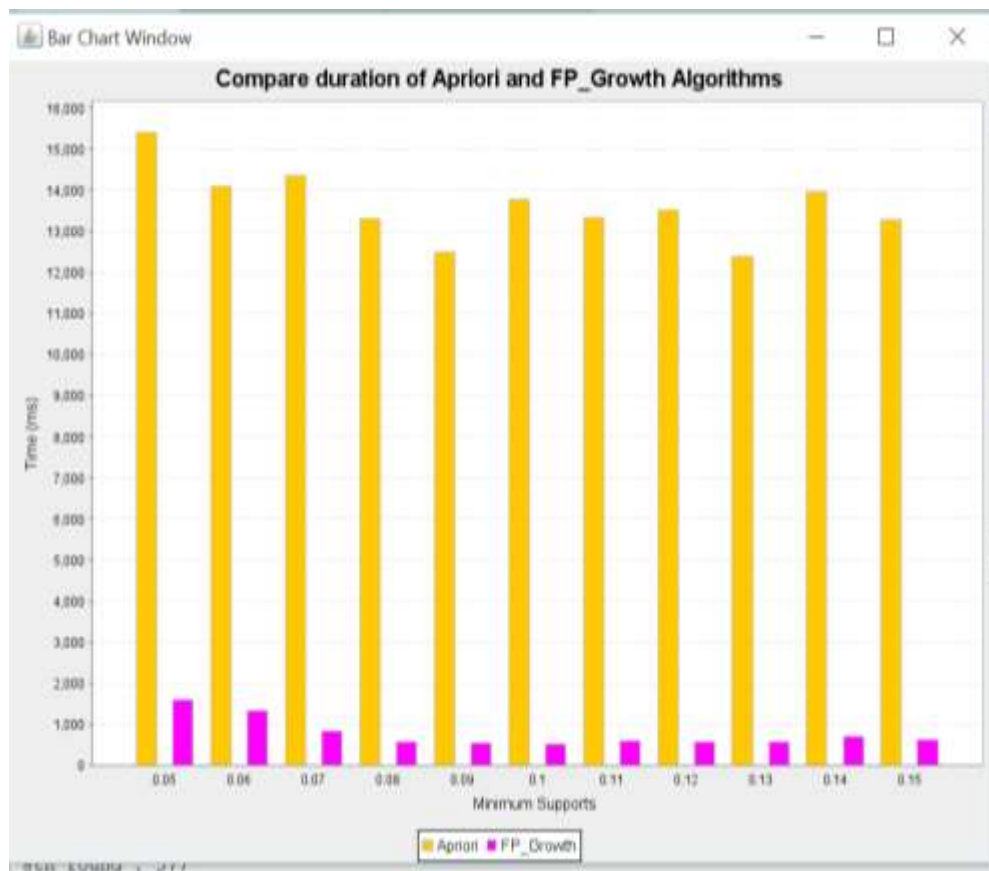
گزارش تمرین سری دوم داده کاوی

فاطمه غلام زاده

۹۵۳۱۰۶۰

سوال ۳ :

برای این سوال مقادیر \min_sup داده شده را در یک آرایه ریخته و سپس در یک حلقه `for` به ازای مقادیر مختلف \min_sup الگوریتم `apriori` و در یک حلقه ی دیگر الگوریتم `FP_Growth` را اجرا نمودم و زمان های متناظر با هر دور را در یک `arraylist` ذخیره کردم. سپس این زمان ها را روی نمودار نشان دادم که به شکل زیر می باشد :



محور افقی مقادیر \min_sup و محور عمودی زمان اجرای الگوریتم ها را بر حسب میلی ثانیه نشان می دهد. همان طور که مشاهده می کنید زمان اجرای الگوریتم `FP_Growth` به ازای مقادیر مختلف \min_sup به مراتب از الگوریتم `apriori` کمتر است.

سوال ۴ :

برای حل این سوال ابتدا باید تمام frequent pattern ها را ایجاد کنیم. سپس آن ها را در حافظه ذخیره کرده ، در یک حلقه به ازای هر frequent pattern چک میکنیم که سایر frequent pattern ها سوپرستی برای آن هستند یا خیر . به این ترتیب می توانیم support را برای سوپرست های هر frequent pattern حساب کنیم و با اعمال شرط چک کنیم که آیا آن frequent pattern ، closed یا max هست یا خیر.

سپس نتایج را در فایل های closed_patterns.txt و max_patterns.txt ذخیره کردم.

تعداد closed ها : ۹۹۸۲

تعداد max ها : ۴۵۵۲

به علت اینکه حجم پروژه برای آپلود زیاد بود فقط فایل های java قرار داده شده اند. کتابخانه های استفاده شده موارد زیر هستند :

gnujaxp.jar

jcommon-1.0.0.jar

jfreechart-1.0.1.jar

junit.jar

remoteExperimentServer.jar

servlet.jar

weka.jar

weka-src.jar

xchart-3.5.4.jar

xchart-demo-3.5.4.jar