



دیتا کاوی - پاییز ۱۳۹۹

نام و نام خانوادگی:
مینو احمدی

تمرین پنجم

گزارش کار

تاریخ:
1399/10/27



دانشکده علوم مهندسی

استاد:
دکتر علی فهیم

مقدمه

همان طور که در صورت پروژه هم گفته شده، در این تمرین قرار است برنامه ای بنویسیم تا تشخیص دهد که یک آیتم ست Derivable است یا Non-Derivable.



نحوه اجرای برنامه

```
python HW5.py itemsets.txt ndi.txt
```

خروجی برنامه

```
> for each itemset in ndi.txt:  
    [1, u] derivable/non-derivable
```

قسمت اول: تابع `get_sup`

این تابع یک `itemset` می گیرد و مقدار `support` آن را بر می گرداند. ورودی اول این تابع یک `set` از `itemset` است و ورودی دوم آن مجموعه آیتیم ست ها و `support` مربوط به آن ها در قالب یک `dict`. خروجی این تابع یک عدد می باشد.

قسمت دوم: تابع `der_nonder`

این تابع که قسمت اصلی کد هم می باشد ورودی هایی همانند تابع `get_sup` دارد. این تابع طبق روشی که در کتاب ذکر شده مجموعه `upper_bound` و `lower_bound` را حساب می کند و با به دست آورد `max_lower_bound` و `min_upper_bound` تشخیص می دهد که آیا `itemset` ورودی `derivable` می باشد یا خیر.

قسمت سوم: تابع `Main`

در `Main` برنامه صرفا کافی است دیتای ورودی را لود کرده و در یک `dict` ذخیره کنیم و سپس به ازای هر `itemset` مدنظر تابع `der_nonder` را اجرا کنیم.
در نهایت خروجی برنامه را چاپ می کنیم.