

ایده ها و توضیحاتی در باره پروژه درس MDA2022

- ✓ داده ها را حداقل بصورت جدولی از زمان ثبت پلاک، پلاک خودرو، شناسه دوربین در نظر بگیرید. برای تحلیل مناسب باید محاسبات در روزهای مختلف تجدید گردد و نتایج تحلیل و تفسیر شود. کدنویسی بصورت comment گذاری شده و گزارش در قالب MS Word در تحویل نهایی لازم است.
- ✓ از روشهای داده کاوی حتما استفاده شود. از یادگیری ماشینی و بهینه سازی برای بهبود نتایج استفاده نمایید.
- ✓ ارزیابی پروژه ها بر مبنای خلاقیت، ارائه تحلیل و نمایش مناسب نتایج در گزارش و کدنویسی روشن و اصولی است. ایده های زیر تنها پیشنهاداتی در این زمینه هستند. ایده های موفق که ذکر نشده اند تشویق می شود.

نمونه ایده ها:

- هیستوگرام هایی از تعداد خودرو ها در تردد های مسیر هر جفت دوربین مجاور (در بازه های از 15 دقیقه تا چند ساعت) که زمان های سفر مختلف از یک دوربین به دوربین دیگر را در روزهای مختلف هفته نشان دهد. دو دوربین مجاورند اگر برای رسیدن از یکی به دیگری دوربین دیگری در مسیر نباشد. مسیرهای با شرایط عادی و غیر عادی در ساعت ها یا روزهای هفته را تحلیل نمایید. در مورد رفتارهای کلی تردها چه میتوانید بگویید؟
- تحلیل های item-basket برای frequent-items و associative rules مثلاً در یافتن مسیرهای مکرر
- با تعریف برداری از تعداد تردد ها در هر ساعت و هر روز هفته (24×7 بعدی) برای هر دوربین روش CF را برای تشخیص شباهت دوربین ها به کار برید. با تقسیم روز به 4 بازه 6 ساعته می توانید بردار کوچکتری هم برای این مقایسه در نظر بگیرید و نتایج را مقایسه کنید. به ایده های مختلف الگوریتم های توصیه گر دیگر یا توابع فاصله متنوع توجه کنید.
- با تعیین خودرو ها در یک بخش گراف و دوربین ها در بخش دیگر روشی مشابه Pixie در تحلیل شباهت دوربین ها و خودرو ها در ساعات مختلف روز، و روزهای مختلف هفته تعیین کنید. روی امکان Query بر این گراف را ایده مطرح کنید.
- از روش HITS برای لیست ارزیابی مهمترین hubs و authorities استفاده کنید.
- ماتریس Utility را بر اساس دوربین ها و خودروهای پرتردد تشکیل دهید و بر اساس SVD، فاکتورهای پنهان این ماتریس را تفسیر نمایید.

موفق باشید