

به نام خدا

هدف پروژه: هدف کلی از این پروژه ، آماده سازی داده هایی است که در پیشبینی آسیب حاد کلیوی در بیمارانی که تحت عمل جراحی قلب قرار می گیرند ، به کار می رود.

در راستای انجام پروژه ، لازم است تا با بررسی مجموعه ی دادگان (ستون های Cr-7 , Cr-1 , Cr) بر اساس روش زیر ، اطلاعات مربوط به هر بیماری به یکی از ۴ برچسب آسیب سطح ۱ ، آسیب سطح ۲ ، آسیب سطح ۳ و فاقد آسیب دسته بندی شود. بر اساس ارزیابی اطلاعات ستون های Cr و Cr-1 یک برچسب برای آسیب روز اول و بر اساس اطلاعات ستون های Cr و Cr-7 نیز یک برچسب بر اساس آسیب روز هفتم برای هر بیمار باید تولید شود.

یعنی هر بیمار دو برچسب دارد ، آسیب روز اول و آسیب روز هفتم پس از عمل جراحی.

- برچسب آسیب سطح ۱: اگر ستون های Cr-1 یا Cr-7 به اندازه ی ۰,۳ از Cr بیشتر یا ۱,۵ تا ۱,۹ برابر Cr شده باشند.
- برچسب آسیب سطح ۲: اگر ستون های Cr-1 یا Cr-7 ۲ تا ۲,۹ برابر Cr باشند.
- برچسب آسیب سطح ۳: اگر ستون های Cr-1 یا Cr-7 مقداری بیش از ۴ داشته باشند یا بیشتر از ۳ برابر Cr باشند.
- برچسب فاقد آسیب: اگر ستون های Cr-1 یا Cr-7 نسبت به Cr افزایش نداشته باشند یا کمتر از 0.3 افزایش یا کمتر از ۱,۵ برابر Cr افزایش داشته باشند.

سوالات پروژه:

با استفاده از روش های حل مساله در کلاس و سایر روش های موجود شروع به آماده سازی مجموعه ی دادگان کنید. این آماده سازی شامل حذف نویز ، حذف مشکل missing value و ساخت برچسب های فوق برای هر داده در مجموعه دادگان می باشد.

- نکته ۱ : به ازای هر داده دو برچسب دارید یکی بر اساس روز اول و دیگری بر اساس روز هفتم پس دو مدل باید آموزش داده شوند.

- نکته ۲: حل مشکلات missing value باید با اتکا به روش های آماری و اعلام شده سر کلاس باشد هم چنین حداقل پنج روش مختلف باید تست و پیاده سازی شود. لازم به ذکر است یکی از این روش ها باید از طریق پیش بینی مقادیر Missing Value از طریق ساخت اعداد رندم مبتنی بر توزیع آماری داده ها باشد.