

به نام خدا
محمد امین قاسمی
گزارش پروژه ۲

LR:

در این مدل هدف مینیمم کردن مجموع قدر مطلق فاصله خط از نقاط از و قید ها هم برای برقراری قدر مطلق اند و چون کمینه سازی نیاز به تعریف متغیر باینری نیست. تابع هدف ۸۶,۸ به دست آمده و بیشترین فاصله یک نقطه از خط هم ۶,۵ شده.

CLR:

این مدل به جای تعریف فقط یک خط، نقطه ها به B دسته تقسیم شده و برای هر کدام خط جدایی در نظر گرفته میشود. شروط 1b و 1c مانند قسمت قبل اند، با این تفاوت که برای نقاط هر دسته اند و در بقیه حالات تفاوتی ایجاد نمیکند. قید 1d هم برای این است که هر نقطه به دقیقاً یک دسته داده شود. با توجه به اینکه بیشترین فاصله قسمت قبل ۶,۵ شده، پس کرانی برای M_i ها است

CLR(maximum absolute difference metric):

مانند CLR است، فقط به جای مجموع فواصل، بیشترین فاصله را کمینه میکند.

CLR(maximum absolute clusterwise difference metric):

مانند CLR است، فقط به جای مجموع فواصل، جمع بیشترین فاصله ها در دسته های مختلف را کمینه میکند.

oCLR:

همانطور که از نتیجه ی CLR پیداست، ممکن است نقطه ای با فاصله ای زیاد از بقیه نقاط یک دسته به آن نسبت داده شده باشد، به خاطر نزدیک بودنش به خط آن دسته.

در حالی که آن نقطه به نقاط دسته ای دیگر نزدیک باشد و ممطقی باشد به آن تعلق گیرد. برای این مشکل در oCLR قیود c_2 تا c_{2e} تضمین میکنند که نقاط یک دسته، مختصات x شان نسبت به سایر نقاط متوالی باشد و به این صورت شرط ترتیبی بر این اساس تعریف میکند.

PWLR:

در این مدل، هدف تعریف یک تابع تکه ای خطی است. در این صورت اگر X_i آخرین نقطه یک دسته باشد، باید نقطه ی r ای بین X_i و X_{i+1} باشد (نقطه برخوردشان) که در معده ی آنها صدق کند که معادله مقابل را نتیجه میدهد.

$$\Rightarrow r = \frac{d_{b+1} - d_b}{c_b - c_{b+1}}$$

$$\Rightarrow X_i \leq \frac{d_{b+1} - d_b}{c_h - c_{h+1}} \leq X_{i+1}.$$

چون با توجه به علامت مخرج، جهت نامساوی عوض میشود، لازم است متغیر هایی تعریف کنیم که نشان دهنده کاهش یا افزایش دسته بعد نسبت به قبل باشد. باینری γ_b برای این منظور است. متغیر های پیوسته $\delta_{i,b}^{+/-}$ بین صفر و یک اگر X_i آخرین نقطه یک دسته باشد، $3d$ و $3e$ تضمین میکنند با توجه به، یکیشان 1 شود و به کمک آنها، بر فرای نامعادله گفته شده را در $3e$ تا $3i$ تضمین میکنیم.

CPWLR:

در این مدل هدف تعرف چند تابع تکه ای خطی برای دسته های مختلف است. پس برای نقاط مرزی لازم به برقراری شرط تکه ای خطی نیست. اگر بخواهیم K دسته و B نقطه شکست داشته باشیم، $K - B$ تکه ی خطی داریم پس باید کران های مربوط به B در شرط ها را تغییر دهیم. همینطور باید متغیر باینری Z_b را متناظر با آخرین یک دسته بودن تعریف کرده و به راست $3c$ و $3d$ اضافه کنیم تا دیگر لازم نباشد پیوستگی رعایت شود. h 4 هم برای این است که K تا دسته داشته باشیم.

(نمودار ها و مقاریر توابع هدف در کد ها هستند)

نوآوری:

به جای شرط نزدیک بودن مختصات x نقاط یک دسته، به تابع هدف مسئله، جمع ماکسیمم فاصله دو نقطه در دسته های مختلف اضافه شده. تغییرات زیر اضافه شده. برای نتیجه بهتر میتوان با توجه به مقیاس های داده ها، جمع وزن دار در نظر گرفت. کد در فایل CLR_max distance in cluster است.

متغیرها: $\gamma_{i,j,b}$ | اگر رابطه در دسته b باشد

$c d_b$: بیشترین فاصله در دسته b

$c d_{-i,j}$: فاصله نقطه i از دسته b

قیدها: $i \neq j$

i) $\delta_{i,b} \in \delta_{j,b} \Rightarrow \gamma_{i,j,b} = 1$

$\gamma_{i,j,b} \geq \delta_{i,b} + \delta_{j,b} - 1$

ii) $c d_b \geq \gamma_{i,j,b} (c d_{-i,j})$

تابع هدف: $\min Z = \left(\sum_{i=1}^n \xi_i \right) w_1 + \left(\sum_{b \in B} c d_b \right) w_2$