

سوال ۱) داریم یک مسئله‌ی طبقه‌بندی دودویی غیرخطی که می‌خواهیم با استفاده از ماشین بردار پشتیبان (SVM) غیرخطی حل کنیم. با این حال، خطاهای طبقه‌بندی برای هر کلاس هزینه‌های متفاوتی دارند. یعنی هزینه‌های خطای نوع اول (FP) و خطای نوع دوم (FN) با هم متفاوت هست، تغییر فرمول SVC به همراه کد پایتون را ارائه دهید.

برای در نظر گرفتن هزینه‌های متفاوت خطاها، می‌توانیم از یک ماتریس وزن‌دهی C استفاده کنیم. این ماتریس، وزن‌های متفاوتی را به نمونه‌های هر کلاس اختصاص می‌دهد. به این ترتیب، SVM تلاش می‌کند تا نمونه‌هایی را که وزن بیشتری دارند، با دقت بیشتری طبقه‌بندی کند.

فرمولاسیون اولیه SVM غیرخطی به صورت زیر است:

$$\text{minimize } 1/2 ||w||^2 + C * \sum(x_i)$$

$$\text{subject to: } y_i * (w^T * \phi(x_i) + b) \geq 1 - x_i, \text{ for all } i$$

در اینجا، w وزن‌های مدل، b بایاس،  $\phi(x_i)$  نگاشت غیرخطی نمونه  $x_i$  به فضای ویژگی‌های با بعد بالاتر،  $y_i$  برچسب کلاس نمونه  $x_i$  و  $x_i$  متغیرهای آزاد هستند.

برای در نظر گرفتن هزینه‌های متفاوت، می‌توانیم فرمولاسیون را به صورت زیر تغییر دهیم:

$$\text{minimize } 1/2 ||w||^2 + \sum(C_i * x_i)$$

$$\text{subject to: } y_i * (w^T * \phi(x_i) + b) \geq 1 - x_i, \text{ for all } i$$

در این فرمولاسیون،  $C_i$  وزن هر نمونه است که با توجه به هزینه‌های خطای مربوط به کلاس آن نمونه تعیین می‌شود.

پیاده‌سازی در پایتون:

برای پیاده‌سازی این فرمولاسیون در پایتون، می‌توانیم از کتابخانه Scikit-learn استفاده کنیم. در این کتابخانه، می‌توانیم از کلاس SVC برای ایجاد یک مدل SVM استفاده کنیم.

```
SVC Python
from sklearn.svm import SVC

# تعریف ماتریس وزن‌دهی C
C_matrix = [[C1, 0], [0, C2]]

# با ماتریس وزن‌دهی SVM ایجاد مدل
clf = SVC(C=C_matrix)

# آموزش مدل
clf.fit(X_train, y_train)
```

در اینجا، C1 و C2 وزن‌های مربوط به کلاس‌های اول و دوم هستند. با تنظیم مناسب این وزن‌ها، می‌توانیم میزان اهمیت هر کلاس را در فرآیند آموزش مدل تعیین کنیم.

### پارامترهای مهم:

- **C:** این پارامتر کنترل می‌کند که چقدر مدل مجاز است خطا کند. مقدار بزرگ‌تر C باعث می‌شود مدل سخت‌گیرانه‌تر شود و کمتر خطا کند.
  - **kernel:** این پارامتر تعیین می‌کند که چه نوع نگاشت غیرخطی برای نمونه‌ها استفاده شود. برخی از گزینه‌های رایج عبارتند از: 'linear', 'poly', 'rbf', 'sigmoid'.
  - **degree:** این پارامتر فقط برای کرنل‌های چندجمله‌ای ('poly') استفاده می‌شود و درجه چندجمله‌ای را مشخص می‌کند.
  - **gamma:** این پارامتر برای کرنل‌های RBF و چندجمله‌ای استفاده می‌شود و میزان تأثیر هر نمونه بر تصمیم‌گیری مدل را کنترل می‌کند.
- با تنظیم مناسب این پارامترها، می‌توان مدل SVM را بهینه کرد و دقت آن را در طبقه‌بندی نمونه‌ها با هزینه‌های متفاوت خطا بهبود بخشید.

### سوال ۲) ارائه ی فرمول ب برای رگرسیون به وسیله ماشین بردار پشتیبان یا همان SVR

رگرسیون بردار پشتیبان (SVR) یک الگوریتم یادگیری ماشین است که برای حل مسائل رگرسیون به کار می‌رود. این الگوریتم بر اساس اصل حداکثر کردن حاشیه (margin) مانند ماشین بردار پشتیبان برای طبقه‌بندی عمل می‌کند، اما با این تفاوت که به جای پیدا کردن یک هایپرپلین برای جداسازی کلاس‌ها، به دنبال یک تابع است که بتواند مقادیر یک متغیر پیوسته را پیش‌بینی کند.

### مفهوم لوله اپس: ILON (ε-tube)

در SVR، مفهوم لوله اپس ILON به کار می‌رود. این لوله یک نوار با عرض  $2\epsilon$  در اطراف تابع رگرسیون است. نقاطی که درون این لوله قرار دارند، به عنوان پیش‌بینی‌های صحیح در نظر گرفته می‌شوند و هیچ جریمه‌ای برای آن‌ها در نظر گرفته نمی‌شود. اما برای نقاطی که خارج از این لوله قرار دارند، یک جریمه خطی متناسب با فاصله آن‌ها از لوله در نظر گرفته می‌شود.

### فرمول بندی ریاضی:

هدف SVR به حداقل رساندن تابع هزینه زیر است:

$$\text{Minimize: } \frac{1}{2} \|w\|^2 + C * \sum (\xi_i + \xi_i^*)$$

Subject to:

$$y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i$$

$$\langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^*$$

$$\xi_i, \xi_i^* \geq 0$$

در این فرمول:

- $w$ : بردار وزن
- $b$ : بایاس
- $x_i$ : بردار ویژگی ورودی
- $y_i$ : مقدار هدف
- $\varepsilon$ : عرض لوله افس ILON
- $\xi_i$  و  $\xi_i^*$ : متغیرهای شل برای حساب کردن خطاهای خارج از لوله

تفسیر:

- جمله اول تابع هزینه، پیچیدگی مدل را به حداقل می‌رساند.
- جمله دوم، خطاهای خارج از لوله افس ILON را جریمه می‌کند.
- قیدها تضمین می‌کنند که مقدار پیش‌بینی شده برای نقاط داخل لوله درون لوله افس ILON قرار داشته باشد و برای نقاط خارج از لوله، خطا جریمه شود.

ترفند هسته:

مانند SVM برای طبقه‌بندی، SVR نیز می‌تواند از ترفند هسته برای نگاشت داده‌ها به یک فضای با بعد بالاتر استفاده کند. این کار به SVR اجازه می‌دهد تا روابط غیرخطی پیچیده بین متغیرهای ورودی و خروجی را مدل کند.