

الف) استفاده از SVM برای مسائل چندکلاسه و تفاوت روش‌های "One-vs-One" و "One-vs-Rest":

SVM به طور ذاتی برای مسائل دوتایی (دوکلاسه) طراحی شده است ولی با این دو روش میتوان مسائل چند کلاسه را هم حل کرد:

One-vs-Rest:

- برای هر کلاس، یک مدل SVM آموزش داده می‌شود که آن کلاس را در مقابل تمام کلاس‌های دیگر تفکیک می‌کند .
 - تعداد مدل‌ها برابر با تعداد کلاس‌ها است .
 - پیش‌بینی نهایی بر اساس بیشترین امتیاز تصمیم‌گیری انجام می‌شود .
- معمولاً برای مجموعه داده‌های کوچک یا زمانی که کلاس‌ها نامتوازن هستند بهتر عمل می‌کند، اما هزینه محاسباتی بالاتری دارد .

One-vs-One

- برای هر جفت کلاس، یک مدل SVM آموزش داده می‌شود .
 - تعداد مدل‌ها برابر با $C(n, 2)$ یا همان ترکیب دو به دو کلاس‌هاست .
 - پیش‌بینی نهایی بر اساس رأی‌گیری اکثریت انجام می‌شود .
- برای مجموعه داده‌های بزرگتر مناسب‌تر است، اما ممکن است در صورت نامتوازن بودن کلاس‌ها دچار مشکل شود .

ب) نقش نقاط پشتیبان (Support Vectors) در تعیین مرز تصمیم‌گیری:

نقاط پشتیبان تنها نمونه‌هایی از داده‌های آموزشی هستند که در نزدیکی مرز تصمیم‌گیری قرار دارند و فاصله آن‌ها از مارجین حداقل است. سایر نقاط داده که خارج از محدوده مارجین هستند، تأثیری بر مرز تصمیم‌گیری ندارند، زیرا حذف آن‌ها تغییری در مرز ایجاد نمی‌کند. اگر همه نقاط در تعیین مرز تصمیم‌گیری دخیل باشند باعث میشود دیتاست‌های متفاوت روی آن تأثیر متفاوتی داشته باشد. از طرفی ما فقط میخواهیم که داده‌ها به درستی دسته‌بندی شوند و این مهم نیست که مثلاً چقدر از مرز فاصله دارند. اگر نقاط مرزی به درستی مشخص شوند نقاط دیگر هم در جای درست قرار می‌گیرند.

ج) تفاوت‌ها و شباهت‌های SVM با Logistic Regression و دلایل ترجیح SVM

شباهت‌ها

- هر دو از الگوریتم‌های یادگیری نظارت‌شده برای مسائل طبقه‌بندی هستند .
- هر دو می‌توانند از کرنل تبدیل غیرخطی برای مدل‌سازی مرزهای پیچیده استفاده کنند. در SVM با تابع کرنل، در Logistic Regression با ویژگی‌های چندجمله‌ای .

تفاوت‌ها

1. تابع هزینه :

- SVM از تابع هینج (Hinge Loss) استفاده می‌کند که بر حداکثر کردن مارجین تمرکز دارد .
- Logistic Regression از تابع Log Loss استفاده می‌کند که بر تخمین احتمالات شرطی تمرکز دارد .

2. حساسیت به داده‌های پرت :

- SVM به دلیل استفاده از مارجین سخت/نرم، نسبت به داده‌های پرت مقاوم‌تر است .
- Logistic Regression می‌تواند تحت تأثیر داده‌های پرت قرار گیرد .

3. کارایی در داده‌های با ابعاد بالا :

- SVM به‌ویژه با کرنل‌های مناسب، در داده‌های با ابعاد بالا عملکرد بهتری دارد .

دلایل ترجیح SVM

- وقتی مرز تصمیم‌گیری غیرخطی است و نیاز به استفاده از کرنل داریم .
- وقتی تعداد ویژگی‌ها بسیار بیشتر از نمونه‌ها است .
- وقتی مقاومت در برابر داده‌های پرت اهمیت دارد .