

تمرین سری اول درس یادگیری عین - رادین خاتم 991572

شکل 1-

الف) اگر داده حذف شده support vector نباشد، مرز تقسیم‌گیری تغییر نمی‌کند. اما اگر support vector باشد، مرز تقسیم‌گیری از داده حذف شده دور می‌شود. در اکثر موارد اجتناب از حذف داده‌ها در تعیین مرز دفعی هستند با حذف یک داده، مرز تقسیم‌گیری تغییر خواهد کرد.

ب) داده‌هایی که غلط طبقه‌بندی شده‌اند  $\epsilon_i > 1$  داده‌هایی که درست طبقه‌بندی شده‌اند  $\epsilon_i \leq 1$  در نظر گرفته می‌شوند.

$$\sum_{i \in \text{misclassified}} \epsilon_i > N \quad \text{number of misclassified delete}$$

$$\sum_{i \in \text{correctly classified}} \epsilon_i > 0$$

$$\boxed{\sum \epsilon_i > N}$$

پ) هر چه قدر  $C$  کوچکتر باشد، اجازه داده می‌شود که داده‌های بیشتری نادرست طبقه‌بندی شوند و سخت‌گیری در تعیین مرز انجام می‌شود. به عنوان مثال اگر  $C=0.5$  مدل ما مثل Hard SVM می‌شود و اگر  $C=100$  به صورت اندک یک مرز در فضا در نظر می‌گیرد.

