



۱. (۴۰ نمره) درست یا نادرست بودن موارد زیر را مشخص کنید:

- داده‌های validation برای تنظیم hyperparameter ها و ارزیابی مدل قبل از آزمایش آن روی داده‌های دیده‌نشده استفاده می‌شود.
- در k-fold cross-validation ، مدل روی k تقسیم مختلف آموزش-تست آموزش داده می‌شود و عملکرد نهایی به عنوان حداکثر امتیاز در میان همه‌ی fold‌ها محاسبه می‌شود.
- هدف اصلی model-selection این است که مدلی انتخاب شود که در مجموعه‌ی آموزشی بهترین عملکرد را داشته باشد.
- feature-selection به بهبود عملکرد مدل کمک می‌کند، زیرا overfitting و پیچیدگی محاسباتی را کاهش می‌دهد.
- دو راهکار رایج برای گسترش رده‌بندهای باینری به طبقه‌بندی چندکلاسه، One-vs-One و One-vs-All (OvA) و (OvO) هستند.
- در k-fold cross-validation ، مقدار k بزرگ‌تر به این معناست که هر بار بخش کوچکتری از داده‌ها برای تست استفاده می‌شود.
- اگر داده‌های آموزشی کم باشند، انتخاب مقدار بزرگ k (مثلاً ۱۰ یا بیشتر) معمولاً بهتر است، زیرا باعث کاهش واریانس می‌شود.
- در حالت ایده‌آل، مقدار k باید برابر با تعداد ویژگی‌های (Features) داده‌ها باشد تا عملکرد مدل بهینه شود.

۲. (۳۰ نمره) در این مسئله، خواهیم دید که چگونه می‌توان یک دسته‌بند را با نگاه کردن به خطاهای آموزش و تست اشکال‌زدایی کرد. در نظر بگیرید که یک دسته‌بند تا همگرایی بر روی مجموعه داده آموزشی D_{train} آموزش داده شده و سپس بر روی مجموعه داده تست D_{test} آزمایش شده است. شما خطای تست را مشاهده می‌کنید و متوجه می‌شوید که مقدار آن بسیار زیاد است. سپس خطای آموزش را محاسبه می‌کنید و متوجه می‌شوید که مقدار آن نزدیک به صفر است.

۱. کدام یک از گزینه‌های زیر انتظار می‌رود که کمک کند؟ تمامی موارد صحیح را انتخاب کنید و انتخاب‌های خود را توضیح دهید.

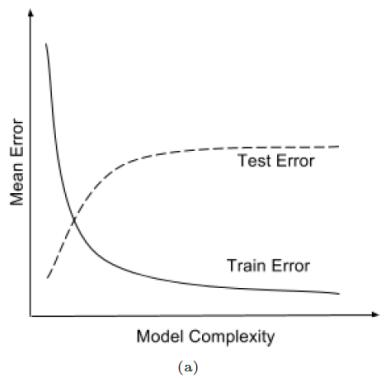
- (آ) افزایش اندازه مجموعه داده آموزشی.
- (ب) کاهش اندازه مجموعه داده آموزشی.
- (ج) کاهش پیچیدگی مدل.

(د) آموزش روی ترکیبی از D_{train} و D_{test} و تست روی

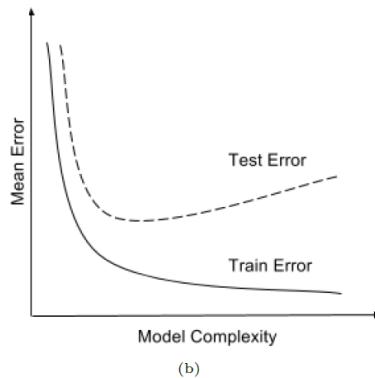
(۵) نتیجه‌گیری که یادگیری ماشین کار نمی‌کند.

۲. این سناریو چه نام دارد؟

۳. فرض کنید که خطای آموزش و تست را به عنوان تابعی از پیچیدگی مدل رسم کرده‌اید. کدام یک از دو نمودار زیر با انتظارات شما همخوانی دارد؟



(a)



(b)

۳. (۳۰ نمره) فرض کنید یک دسته‌بند خطی در فضای ۳-بعدی با وزن‌های اولیه

$$\mathbf{w} = (2, -1, 3), \quad b = -2$$

آموزش دیده است. حال داده‌ی جدیدی به صورت $\mathbf{x} = (x_1, x_2, x_3) = (x_1, x_2, x_3)$ را دریافت می‌کنیم.

الف) اگر مقدار x_3 به شدت افزایش یابد، چگونه روی خروجی مدل و Decision-Boundary تأثیر می‌گذارد

ب) اگر در مرحله‌ی تست، مقدار ویژگی x_1 همیشه مقدار ثابتی (مثلًاً ۵) داشته باشد، در حالی که در آموزش متغیر بوده است، این وضعیت چه تأثیری بر عملکرد مدل خواهد داشت؟ چرا؟