**به نام خدا**

**درس یادگیری ماشین**
**نیم‌سال دوم ۰۳-۰۴**
**استاد: ربیعی- دهقانیان**

**دانشکده مهندسی کامپیوتر**

---

**آزمونک دوم - ۱۴۰۴/۰۱/۱۷** زمان آزمون: ۳۰ دقیقه

---

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.

**درست**

- 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.

**نادرست، عملکرد نهایی به عنوان میانگین همه ی هاfold است.**

- هدف اصلی model-selection این است که مدلی انتخاب شود که در مجموعه‌ی آموزشی بهترین عملکرد را داشته باشد.

**نادرست، در مجموعه ی تست بهترین عملکرد را داشته باشد.**

- feature-selection به بهبود عملکرد مدل کمک می‌کند، زیرا overfitting و پیچیدگی محاسباتی را کاهش می‌دهد.

**درست**

- دو راهکار رایج برای گسترش رده‌بندهای باینری به طبقه‌بندی چندکلاسه، One-vs-All (OvA) و One-vs-One (OvO) هستند.

**درست**

- در k-fold cross-validation ، مقدار k بزرگ‌تر به این معناست که هر بار بخش کوچکتری از داده‌ها برای تست استفاده می‌شود.

**درست**

- اگر داده‌های آموزشی کم باشند، انتخاب مقدار بزرگ k (مثلاً ۱۰ یا بیشتر) معمولاً بهتر است، زیرا باعث کاهش واریانس می‌شود.

**درست**

- در حالت ایده‌آل، مقدار k باید برابر با تعداد ویژگی‌های (Features) داده‌ها باشد تا عملکرد مدل بهینه شود.

**نادرست، مقدار k ربطی به تعداد ویژگی ندارد.**

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 دادهٔ آموزشی $D_{train}$ آموزش داده شده و سپس بر روی مجموعه دادهٔ 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 انتخاب های خود را توضیح دهید.

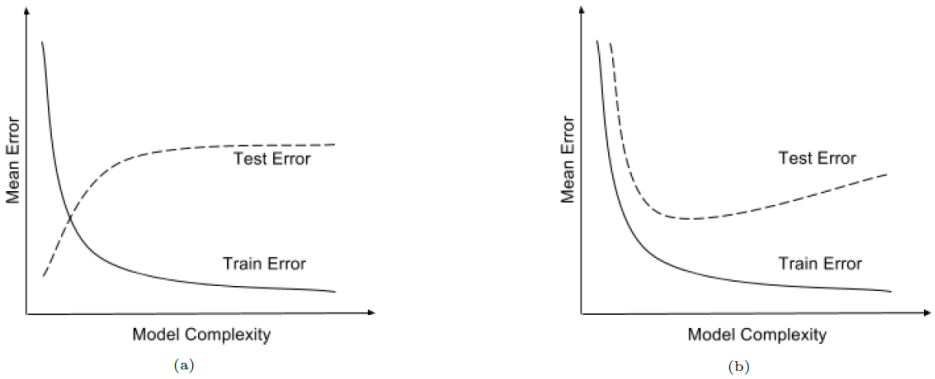
(آ) افزایش اندازهٔ مجموعه دادهٔ آموزشی.
(ب) کاهش اندازهٔ مجموعه دادهٔ آموزشی.
(ج) کاهش پیچیدگی مدل.
(د) آموزش روی ترکیبی از $D_{train}$ و $D_{test}$ و تست روی $D_{test}$.
(ه) نتیجه‌گیری که یادگیری ماشین کار نمی‌کند.

**گزینه ی آ و ج. چون مدل دارد overfit میشود یا باید تعداد داده های train را زیاد کنیم یا پیچیدگی مدل را کاهش دهیم.**

۲. این سناریو چه نام دارد؟

**overfit**

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؟




**نمودار b ، چون هنگامی که پیچیدگی مدل زیاد میشود خطای اموزش کم می شود اما خطای تست از یک جا به بعد زیاد میشود.**

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

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

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

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

مدل ما از نوع خطی است و تابع تصمیم آن به صورت زیر تعریف می‌شود:

$$f(\mathbf{x}) = 2x_1 - x_2 + 3x_3 - 2$$

**تأثیر بر خروجی مدل:** اگر مقدار $x_3$ به شدت افزایش یابد (مثلاً خیلی بزرگ شود)، آنگاه جمله‌ی $3x_3$ در تابع تصمیم مقدار بسیار بزرگی خواهد شد. در نتیجه خواهیم داشت:

$$f(\mathbf{x}) \to +\infty$$

بنابراین، خروجی مدل بسیار مثبت خواهد شد و مدل با اطمینان بالا داده را در کلاس مثبت قرار خواهد داد.

**تأثیر بر Boundary Decision (مرز تصمیم):** مرز تصمیم با استفاده از معادله زیر تعریف می‌شود:

$$2x_1 - x_2 + 3x_3 - 2 = 0$$

این یک صفحه‌ی ثابت در فضای سه‌بعدی است که فقط به **وزن‌ها و بایاس** وابسته است، نه به ورودی‌ها. چون وزن‌ها و بایاس تغییر نکرده‌اند، مرز تصمیم نیز **تغییری نمی‌کند**.

**نتیجه‌گیری:** افزایش شدید $x_3$ موجب افزایش خروجی مدل می‌شود، اما تأثیری بر موقعیت یا جهت مرز تصمیم ندارد.

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

در زمان آموزش، وزن $w_1 = 2$ برای ویژگی $x_1$ آموخته شده است، که نشان می‌دهد $x_1$ در داده‌های آموزشی متغیر بوده و در تصمیم‌گیری مدل نقش داشته است.

اگر در مرحله‌ی تست، مقدار $x_1$ همیشه ثابت شود (مثلاً برابر با ۵)، آنگاه این ویژگی دیگر کمکی به تمایز بین کلاس‌ها نمی‌کند. با این حال، مدل همچنان وزن $w_1$ را در تصمیم‌گیری لحاظ می‌کند، زیرا انتظار دارد $x_1$ مانند آموزش، متغیر و دارای اطلاعات باشد.

این اختلاف بین توزیع داده‌های آموزش و تست (که به آن *distribution-shift* گفته می‌شود)، می‌تواند باعث کاهش دقت مدل در زمان تست شود. زیرا مدل در حال استفاده از اطلاعاتی است که دیگر کاربردی ندارند.

**نتیجه‌گیری:** ثابت بودن $x_1$ در مرحله‌ی تست، در حالی که در آموزش متغیر بوده، موجب کاهش دقت مدل می‌شود، زیرا یکی از منابع اطلاعاتی مؤثر در مدل‌سازی از بین رفته است.