

به نام خدا

دانشگاه صنعتی شریف - دانشکده مهندسی کامپیوتر

مقدمه‌ای بر یادگیری ماشین

## فصل سوم: یادگیری، ارزیابی و تنظیم کردن مدل‌ها

### مبحث: اعتبارسنجی متقابل (Cross (Validation

نویسنده: پیمان ناصری

#### هدف

در یادگیری ماشین، قابلیت تعمیم به توانایی یک الگوریتم در تأثیرگذاری روی ورودی‌های مختلف اشاره دارد؛ به این معنی که مدل یادگیری ماشین توانایی عملکرد خود را در برابر ورود داده‌های جدید که در داده‌های آموزشی نبوده از دست نمی‌دهد.

برای انسان، عمومیت بخشی و تعمیم، طبع ترین کار ممکن است. ما می‌توانیم به صورت آبی (در ذهنمان) دسته بندی و طبقه بندی انجام دهیم. به طور مثال، ما بدون شک یک سگ را تشخیص می‌دهیم حتی اگر قبل از آن، چنین نژادی را ندیده باشیم. با این حال چنین مساله ای می‌تواند برای یک مدل

یادگیری ماشین کاملاً چالش برانگیز باشد. به همین دلیل بررسی توانایی الگوریتم در عمومیت بخشی کار مهمی است که در زمان ساخت مدل نیاز به توجه زیادی دارد. برای انجام چنین کاری از روش اعتبارسنجی متقابل (CV) استفاده می کنیم.

## اعتبارسنجی متقابل (Cross Validation)

---

این روش سعی می کند یک تخمین نااریب (unbiased) از generalization error بزند. برای مدل کردن اعتبارسنجی متقابل روشهای متنوع زیادی ممکن است استفاده

شود با اینکه همه ی آنها یک الگوریتم مشابه دارند

1- تقسیم کردن مجموعه دادهها به دو بخش: بخشی برای آموزش و بخش دیگر برای ارزیابی

2- مدل را روی مجموعه دادههای آموزشی، آموزش دهید

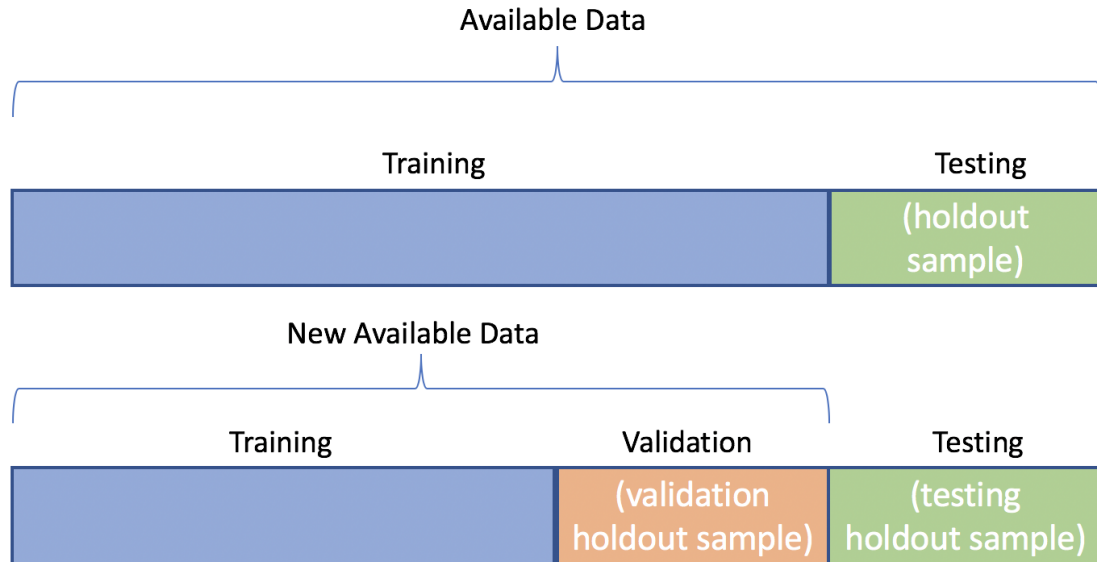
3- اعتبار مدل را روی داده های ارزیابی بدست میاوریم

4-مرحله یک تا سه را چندین بار تکرار میکنیم . تعداد تکرار ها بستگی روشی دارد که ما برای CVمورد استفاده قرار میدهیم .

## انواع روش های اعتبارسنجی

---

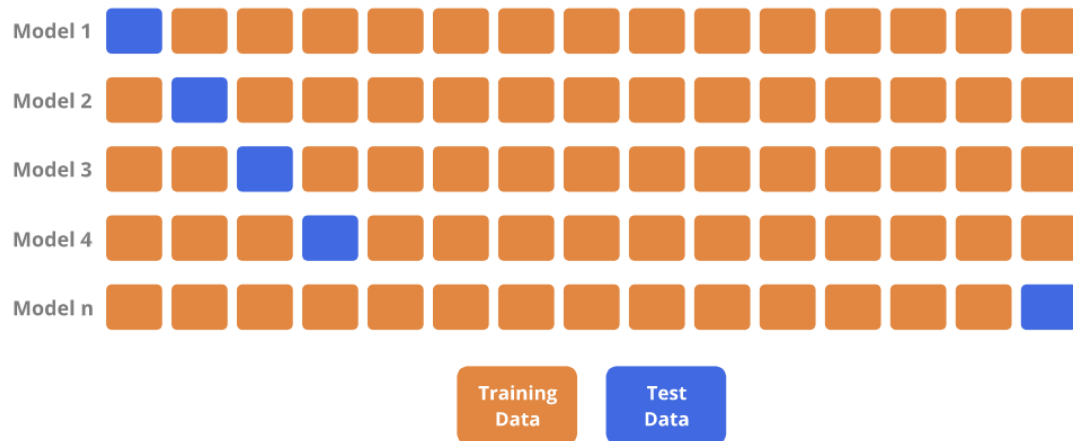
## Holdout Validation



در این روش، داده‌ها به دو دسته train و test تقسیم می‌شوند. این تقسیم می‌تواند به صورت 40/60 ، 30/70 یا 20/80 باشد. بنابراین مدل مورد نظر روی داده‌های train آموزش دیده و روی داده‌های test مورد ارزیابی قرار می‌گیرد. معمولاً از روش اعتبارسنجی Hold-out برای مجموعه داده‌های بزرگ استفاده می‌کنیم زیرا تنها به یک بار آموزش مدل نیاز دارد. البته روش Hold-out یک ایراد اساسی دارد. برای مثال، مجموعه داده‌ای که توزیع کاملاً یکنواختی ندارد، در چنین حالتی ممکن است پس از تقسیم، در یک نقطه‌ی نقطه‌ی ناهموار قرار بگیریم. به‌طور مثال مجموعه‌ی آموزشی، مجموعه تست را نشان نخواهد داد. ممکن است مجموعه آموزشی و تست تفاوت بسیاری داشته باشند، ممکن است یکی از آنها سخت‌تر یا آسان‌تر باشد. به‌علاوه این واقعیت که ما مدلمان را فقط یک بار تست می‌کنیم ممکن است یک تنگنا برای این روش باشد. به دلایلی که اشاره شد، ممکن است نتایج به‌دست آمده به روش Hold-out نادرست به نظر برسد.

## (LOOCV (Leave One Out Cross Validation

## Leave-One-Out Cross Validation



در این تکنیک، از تمامی داده‌ها به جز یک داده برای آموزش و از داده باقی مانده برای تست مدل استفاده می‌شود. این فرآیند  $N$  بار تکرار می‌شود که  $N$  تعداد داده‌ها را نشان می‌دهد. مزیت این روش این است که از تمامی داده‌ها برای آموزش و تست مدل استفاده خواهد شد. البته از نظر محاسباتی روش پرهزینه‌ای است چراکه زمان زیادی برای اعتبارسنجی متقابل مدل با استفاده از LOOCV نیاز است.

## K-Fold Cross Validation



در این روش، داده‌ها به  $k$  بخش تبدیل می‌شوند. هرکدام از داده‌ها به صورت

تصادفی در یکی از این  $k$  بخش قرار می‌گیرند. در این روش، آموزش و تست  $k$  بار تکرار می‌شوند. در هر تکرار  $k-1$  بخش به عنوان داده train و یکی به عنوان test در نظر گرفته می‌شود. خطای مدل برابر با میانگین مدل در  $k$  تکرار است. این روش در واقع تلاش می‌کند معایب دو روش بالا را کم کند و از مزی