

تحقیق جلسه ۲

یادگیری ماشین ۳

مسعود شیری

## ترکیب مستقل خطی

دو تابع  $g(x)$ ،  $f(x)$  را وابسته خطی میگوییم هرگاه عدد ثابتی مانند  $k \in \mathbb{R}$  وجود داشته باشد به طوری که  $f(x) = k \cdot g(x)$ . به بیان ساده تر، دو تابع را وابسته خطی میگوییم اگر یکی مضرب ثابتی از دیگری باشد. دو تابع  $f$  و  $g$  که وابسته خطی نیستند را مستقل خطی می‌نامیم. مثال:

دو تابع  $f(x) = x^2 + x$  و  $g(x) = 2x^2 + 2x$  وابسته خطی اند و توابع  $f(x) = x^2 + x$  و  $h(x) = 4x^2 + 2x$  مستقل خطی اند. [1]

## Validation Data Set

تمامی داده‌هایی که برای ایجاد یا تست مدل‌های یادگیری ماشینی استفاده می‌شود، مجموعه داده (Data Set) نامیده می‌شود. مجموعه داده‌ها به ۳ گروه تقسیم می‌شوند: Training data, Validation data, Test data

Validation data (یا داده‌های اعتبارسنجی) برای تنظیم پارامترهای مدل و مقایسه مدل‌های مختلف به کار می‌روند. هدف مقایسه مدل‌های مختلف و تعیین بهترین مدل می‌باشد. داده‌های اعتبارسنجی باید با داده‌های آموزشی تفاوت داشته باشند. از داده‌های اعتبارسنجی نباید در مرحله آموزش استفاده شوند. در غیر این صورت، این مدل بیش برآزش (یا Over fit) می‌شود. در نتیجه تعمیم آن به اطلاعات جدید (تولید شده) ضعیف است. [2]

## MSE loss

یکی از معروف‌ترین توابع زیان در تحلیل رگرسیونی، خطای میانگین مجزور (Mean Squared Error) است که به اختصار MSE نامیده می‌شود. این تابع زیان، میانگین مربعات فاصله بین مقدار پیش‌بینی و واقعی را محاسبه می‌کند. شیوه و با استفاده از فرمول زیر بدست می‌آید:

$$MSE = \frac{\sum (y_i - \hat{y}_i)^2}{n}$$

در مباحث آماری، معمولاً به چنین تابعی، زیان  $L2$  گفته می‌شود. با توجه به استفاده از توان ۲ در محاسبه  $MSE$ ، شکل این تابع زیان برحسب مقادارهای پیش‌بینی (یا خطا) به صورت سهمی خواهد بود. [3]

## Gradient Descent

گرادیان کاهشی (Gradient Descent) یک الگوریتم بهینه‌سازی تکراری مرتبه اول برای یافتن کمینه‌های موضعی یا Local Minimum روی یک تابع مشتق‌پذیر است.

برای یافتن کمینه‌های یک تابع یا به طور کلی برای یافتن اکسترمم‌های یک تابع (کمینه یا بیشینه) می‌توانیم مشتق آن را بگیریم و مساوی صفر قرار دهیم. جواب‌های چنین معادله‌ای نقاطی از تابع را خواهند داد که تابع مقدار minimum یا maximum (کمینه یا بیشینه) دارد.

سوال اینجاست که اگر یافتن کمینه یک تابع به این سادگی است چرا باید از روش گرادیان کاهشی استفاده کرد؟ جواب این است که گاه توابع چنان پیچیده هستند (مثلاً چند متغیره یا تک متغیره از مرتبه‌های بالا) که محاسبه مشتق آن‌ها و حل معادله آن‌ها به سادگی ممکن نیست.

در چنین شرایطی باید سراغ روش‌های ابتکاری برویم. روش‌هایی که معمولاً سعی می‌کنند با یک جستجوی هوشمندانه تکراری و تصحیح مداوم خود در یک مدت زمان معقول به جواب‌های معقول به تعداد کافی برسند. شاید این جواب‌ها جواب قطعی نباشند اما به اندازه کافی خوب هستند و در اغلب مسائل همین کافی است. [4]

منابع:

1. <https://irmath.com/index.php/learning/differentioalequations/linear-equations/402-linearly-dependent-and-independent-functions>
2. <https://www.nemoudar.com/blog/8-machine-learning-expressions>
3. <https://blog.faradars.org/loss-function-in-machine-learning>
4. <https://hamrueyesh.com/what-is-a-decreasing-gradient>