



خلاصه دوره‌های هوش مصنوعی گوگل – یادداشت مطالعه به فارسی

دوره فشرده یادگیری ماشین

Machine Learning Crash Course

ارائه‌دهنده: Google for Developers

۱۲ ماژول با تمرین تعاملی

سطح: میانی – نیازمند پایتون و ریاضیات پایه

فایل ۱۵ از ۲۲ – تهیه شده برای تدریس

موضوع: مبانی یادگیری ماشین از رگرسیون تا شبکه عصبی و LLM **تعداد درس:** ۱۲ ماژول **سطح:** میانی

پیش‌نیاز: پایتون، جبر و آمار مقدماتی **قالب:** متن + ویدئو + تمرین تعاملی

خلاصه کلی دوره

قدیمی‌ترین و جدی‌ترین دوره فنی گوگل در این فهرست، که چند سال پیش کاملاً بازنویسی شد و حالا یک ماژول مستقل درباره مدل‌های زبانی بزرگ هم دارد. برخلاف دوره‌های Gen AI که مصرف‌کننده‌محورند، این دوره سازنده‌محور است: می‌خواهد بفهمید مدل واقعاً چطور یاد می‌گیرد. دوره متن‌محور با ویدئوهای کوتاه و تمرین‌های تعاملی داخل مرورگر است. توصیه گوگل این است که مبتدی‌ها ماژول‌ها را به ترتیب بگذرانند.

نکات مهم هر ماژول

بخش اول – مدل‌های یادگیری ماشین

۱. رگرسیون خطی

- ساده‌ترین مدل: پیش‌بینی یک عدد پیوسته.
- مفاهیم پایه‌ای که در کل ML تکرار می‌شوند: مدل خطی، تابع زیان (loss)، گرادیان کاهشی (gradient descent)، و تنظیم فرایارامترها.
- فهم عمیق این ماژول، بقیه دوره را آسان می‌کند.

۲. رگرسیون لجستیک

- مدلی که به جای عدد، احتمال پیش‌بینی می‌کند.
- پایه اکثر مسائل طبقه‌بندی.

۳. طبقه‌بندی

- طبقه‌بندی دودویی و انتخاب آستانه (threshold).
- ماتریس درهم‌ریختگی (confusion matrix) و معیارها: دقت، صحت (precision)، بازخوانی (recall)، AUC.
- نکته کلیدی: «دقت بالا» به تنهایی معیار بی‌معنایی است وقتی داده نامتوازن باشد.

بخش دوم – داده

۴. کار با داده عددی

- تحلیل و تبدیل ویژگی‌ها: نرمال‌سازی، مقیاس‌بندی، مدیریت مقادیر پرت.

۵. کار با داده رشته‌ای

- one-hot encoding, feature hashing, mean encoding

- چالش ویژگی‌های با کاردینالیتی بالا.

۶. مجموعه داده، تعمیم و بیش‌برازش

- تقسیم داده به آموزش/اعتبارسنجی/آزمون.
- بیش‌برازش (overfitting): مدل داده را حفظ کرده، نه یاد گرفته.
- تعمیم (generalization) هدف نهایی است، نه عملکرد روی داده آموزش.

بخش سوم – مدل‌های پیشرفته

۷. شبکه‌های عصبی

- پرسپترون، لایه‌های پنهان، و توابع فعال‌سازی.
- چرا افزودن لایه، توانایی مدل‌سازی روابط غیرخطی را می‌دهد.

۸. تعبیه‌ها (Embeddings)

- نمایش داده‌های با ابعاد بالا در فضای برداری کوچک‌تر و معنادار.
- مبنای جست‌وجوی معنایی، سامانه‌های توصیه‌گر و بازیابی در RAG.

۹. آشنایی با مدل‌های زبانی بزرگ

- مسیر «از توکن تا ترنسفورمر».
- چگونه مدل یاد می‌گیرد متن بعدی را پیش‌بینی کند.
- پل ارتباطی این دوره کلاسیک با دنیای Gen AI.

بخش چهارم – ML در دنیای واقعی

۱۰. سامانه‌های ML در محیط عملیاتی

- اجزای یک سامانه واقعی: خط لوله داده، آموزش، سرو کردن، پایش.

- واقعیت مهم: کد مدل، بخش کوچکی از سامانه است.

۱۱. AutoML

- خودکارسازی انتخاب مدل و تنظیم فرآپارامتر؛ کجا مفید است و کجا نه.

۱۲. انصاف در ML

- شناسایی انواع **سوگیری**: سوگیری گزارش، سوگیری انتخاب، سوگیری خودکارسازی.
- روش‌های سنجش و کاهش نابرابری در عملکرد مدل روی گروه‌های مختلف.

جمع‌بندی سریع برای مرور

ستون فقرات دوره

مدل ← تابع زیان ← گرادیان کاهشی ← ارزیابی ← تعمیم. هر مدل پیشرفته‌ای فقط نسخه پیچیده‌تری از همین چرخه است.

- اگر وقت محدود دارید: ماژول‌های ۱، ۳، ۶ و ۹ بیشترین بازده را دارند.
- ماژول انصاف (۱۲) را به عنوان «ضمیمه اخلاقی» ببینید؛ در عمل تعیین می‌کند مدل شما قابل استقرار هست یا نه.
- این دوره پیش‌نیاز فکری خوبی برای مسیر پیشرفته Gen AI برای توسعه‌دهندگان است.

منابع

• <https://developers.google.com/machine-learning/crash-course>

• [/https://ai.google/learn-ai-skills](https://ai.google/learn-ai-skills)