



خلاصه دوره‌های هوش مصنوعی گوگل – یادداشت مطالعه به فارسی

آشنایی با هوش مصنوعی مسئولانه

Introduction to Responsible AI

ارائه‌دهنده: Google Cloud / Google Skills

دوره ریزآموزشی – حدود ۱ ساعت

سطح: مقدماتی – رایگان با نشان

فایل ۱۰ از ۲۲ – تهیه شده برای تدریس

موضوع: هفت اصل هوش مصنوعی گوگل و معنای عملی آن‌ها **تعداد درس:** دوره ریزآموزشی **سطح:** مقدماتی
مدت: حدود ۱ ساعت **خروجی:** نشان دیجیتال

خلاصه کلی دوره

سومین ضلع مثلث مقدماتی گوگل (همراه با «آشنایی با AI مولد» و «آشنایی با LLM»). این دوره توضیح می‌دهد چرا هوش مصنوعی مسئله‌ای یک بحث اخلاقی انتزاعی نیست، بلکه یک الزام عملی برای هر کسی است که سامانه AI می‌سازد یا استفاده می‌کند.

محور دوره **اصول هوش مصنوعی گوگل** است و اینکه این اصول چطور به تصمیم‌های روزمره محصول ترجمه می‌شوند.

نکات مهم هر بخش

۱. چرا مسئولیت‌پذیری اهمیت دارد

- AI بی‌طرف نیست: داده‌ای که از دنیای واقعی می‌آید، نابرابری‌های همان دنیا را با خود می‌آورد.
- تصمیم‌های به ظاهر فنی (انتخاب داده، تعریف معیار موفقیت) در واقع **تصمیم‌های ارزشیاند**.
- هزینه بی‌مسئولیتی: آسیب به کاربر، از دست رفتن اعتماد، ریسک قانونی و برندی.

۲. اصول هوش مصنوعی گوگل

- برای جامعه **سودمند** باشد.
- از ایجاد یا تقویت **سوگیری ناعادلانه** پرهیز کند.
- برای **ایمنی** ساخته و آزموده شود.
- در برابر انسان **پاسخگو** باشد.
- اصول **حریم خصوصی** را رعایت کند.
- استانداردهای بالای **تعالی علمی** را حفظ کند.
- تنها برای کاربردهایی در دسترس قرار گیرد که با این اصول هم‌خوان‌اند.

۳. کاربردهایی که گوگل دنبال نمی‌کند

- فناوری‌هایی که آسیب کلی‌شان بر منفعتشان می‌چربد.

- سلاح یا فناوری‌هایی که هدف اصلی‌شان آسیب رساندن به انسان است.
- سامانه‌های نظارتی که هنجارهای پذیرفته‌شده بین‌المللی را نقض می‌کنند.
- کاربردهای مغایر با حقوق بین‌الملل و حقوق بشر.

۴. از اصل تا عمل

- اصول به‌تنهایی کافی نیستند؛ باید به **رویه** تبدیل شوند: بازبینی پیش از انتشار، سنجش عملکرد روی گروه‌های مختلف، مسیر گزارش مشکل.
- **مسئولیت مشترک** است: مهندس، طراح، مدیر محصول و کاربر نهایی همه سهم دارند.
- ابزارها کمک می‌کنند اما جای قضاوت انسانی را نمی‌گیرند.

۵. تصمیم‌گیری در شرایط مبهم

- بیشتر موارد واقعی «درست/غلط» روشن ندارند؛ مبادله (trade-off) دارند.
- روش عملی: **شناسایی زودهنگام مسئله** (issue spotting) – پیش از ساخت بپرسید چه کسی ممکن است آسیب ببیند.
- مستندسازی تصمیم و دلیلش، تا بعداً قابل بازبینی باشد.

جمع‌بندی سریع برای مرور

هسته دوره

هوش مصنوعی مسئولانه یعنی از **قبل** بپرسیم «این ممکن است به چه کسی و چطور آسیب بزند؟» – نه اینکه بعد از حادثه توضیح بدهیم.

- هفت اصل را حفظ نکنید؛ یاد بگیرید هرکدام در تصمیم مشخصی چه معنایی می‌دهد.
- سوگیری کاملاً حذف نمی‌شود؛ **سنجیده و مدیریت** می‌شود.
- گام بعدی عملی: دوره «Responsible AI: Applying AI Principles with Google Cloud» که همین اصول را به رویه سازمانی تبدیل می‌کند.

منابع

- https://www.skills.google/course_templates/554

/https://ai.google/responsibility/principles •

https://cloud.google.com/blog/topics/training-certifications/new-generative-ai-trainings-from-google-cloud •