

رفتار واقعی هوش مصنوعی برای دانشمندان داده

انتشارات

تاریخ انتشار

فهرست مطالب

ii	تشکر نامه
iii	فهرست همکاران و مشارکت کنندگان
۱	۱ مقدمه: ماشین‌های اخلاقی
۱	۱.۱ ماشین‌های اخلاقی
۳	۲.۱ علم داده چیست؟
۴	۳.۱ موارد مطالعاتی
۴	۱.۳.۱ مورد اول اخلاق تحقیق و روش علمی
۵	۲.۳.۱ زیر بخش اول
۵	۴.۱ بخش دوم

تشکر نامه

مایلم از مشارکت کنندگان زیر برای کمک‌های عالی و متنوعشان در این کتاب تشکر کنیم: پیتر هرشوگ (اخلاق بودایی)، جان هکر رایت (اخلاق فضیلت)، ساموئل جی لوین و دانیل سینکلر (اخلاق یهودی)، کالین مارشال (اخلاق دئونتولوژیک)، جوی میلر و آندریا سالیوان کلارک (اخلاق بومی) و جان مورانگی (اخلاق آفریقایی). هدف ما ترسیم تصویری اخلاقی است که تا حد امکان متنوع و جذاب باشد. بدون کمک این عزیزان خردمند، توانستیم این کتاب را ایجاد کنیم!

فهرست همکاران و مشارکت کنندگان

همکاران:

پیتر سینگر

مرکز دانشگاهی برای ارزش های انسانی
دانشگاه پرینستون

بیپ فای تسه

مرکز دانشگاهی برای ارزش های انسانی
دانشگاه پرینستون

سامول جی لوین

مرکز حقوقی تورو

کولین مارشال

گروه فلسفه دانشگاه واشنگتن

جویی میلر

گروه فلسفه دانشگاه وست چستر

جان مورونگی

گروه فلسفه دانشگاه وست چستر

مشارکت کنندگان:

جان هکر رایت

گروه فلسفه دانشگاه گوئلف

دنیل سینکلر

دانشکده حقوق دانشگاه فوردهام

پیتر دی هرشوک

مرکز شرقی-غربی

آندریا سالیوان کلارک

گروه فلسفه دانشگاه ویندزور

فصل ۱

مقدمه: ماشین‌های اخلاقی

۱.۱ ماشین‌های اخلاقی

این کتاب، برای دانشمندان داده و افراد علاقه‌مند به این حوزه است؛ که در برخی جهات آن دچار تردید شده‌اند. یا به طور خلاصه، راه درست و غلط استفاده از دیتا را به افراد نشان دهد و از استفاده‌ی غیراخلاقی آن جلوگیری کند.

به نظر می‌رسد که اهمیت علم داده در زندگی روزمره بسیار کم است! این امر باعث می‌شود که مردم عادی حتی در درک کردن این فناوری قدرتمند، کاملاً ناتوان باشند؛ چه رسد به شکل‌دهی یا اداره‌ی آن! از طرفی خیلی از دانشمندانی که در این زمینه مشغول فعالیت هستند، نه زمان کافی برای کسب معلومات اخلاقی را دارند و نه منابع کافی برای اینکه ذهن خود را در گیر اهمیت اخلاق در این زمینه کنند. در صورتی که این فناوری می‌تواند تأثیرات اخلاقی زیادی را بر جامعه وارد کند. این کتاب در جهت کاهش این کمبودها نوشته شده است تا چراغ راهی باشد برای کسانی که به اخلاق در این حوزه اهمیت می‌دهند. برای این ایده که «دانشمندان باید اخلاق را بیاموزند»، تفکراتی مانند «شما نمی‌توانید چیزی در مورد اخلاق به کسی بیاموزید، مردم آن را می‌سازند» وجود دارد. البته قسمتی از آن درست است، مردم یک جامعه، اخلاق را می‌سازند.

امروزه ما با یک پدیده‌ی بسیار قدرتمند و البته بسیار پر خطر به نام «علم داده» روبرو هستیم؛ بنابراین، باید اخلاقیات و ضوابط این حوزه به صورت گسترده آموزش داده شود.

برای اینکه این کتاب تا حد امکان مفید و دوستانه واقع شود، سعی کردیم مطالب را با لحنی ساده بیان کنیم. در این کتاب، ۷ مثال واقعی که استفاده نادرست از علم داده را نشان می‌دهند، بیان می‌کنیم. ما همچنین با چندین دانشمند برجسته‌ی اخلاق تماس گرفتیم تا در هر مورد نظراتشان را بپرسیم. همچنین برای ارائه‌ی طیف وسیع رفتارها و اخلاقیات انسانی، از سه دیدگاه غرب نسبت به اخلاق، فراتر رفتیم، سه رویکرد غرب عبارتند از: نتیجه‌گرایی (فایده‌گرایی)، دین‌شناسی و رفتار با تقوا. رویکردهایی که به طور اضافی بررسی کردیم: بودایی، یهودی، بومی و آفریقایی. هر یک از این رفتارها و رویکردها، می‌توانند زاویه‌ی دید متنوعی را ارائه کنند که ممکن است به آن فکر نکرده باشیم. هدف ما این است که درک کاملی از هر رویکرد ارائه دهیم، یک جعبه ابزار کامل برای روبرویی با چالش‌های آینده.

همانطور که می‌دانیم، یک مشکل خاص، می‌تواند با زوایای دید متفاوت (رویکردهای اخلاقی متفاوت که اشاره کردیم)، به طور مختلف تحلیل و بررسی شود. توانایی تحلیل معضل از دیدگاه‌های مختلف، لازمه‌ی «تفکر انتقادی» است. امیدوارم این کتاب دیدگاه گسترده‌ای را در اختیار خواننده قرار دهد!

۲.۱ علم داده چیست؟

علم داده، اصولی است برای استخراج دیتاهای غیر بدیهی و الگوها از مجموعه دیتاهای بزرگ. از طرفی هوش مصنوعی را می‌توان هر گونه پردازش اطلاعات که کارکرد روانی را انجام می‌دهد، اطلاق کرد. مثلا پیش‌بینی، تداعی کردن، تخیل کردن، برنامه‌ریزی و به طور کلی، هر پردازشی که تا کنون موجودات زنده قادر به انجام آن بودند.

ماشین لرنینگ (ML)، زیرمجموعه‌ای از علم داده و بخش رو به رشدی از این زمینه است. بر خلاف GOFAI، ماشین لرنینگ (ML) شکلی از هوش مصنوعی است که از رویکردهای آماری برای یافتن الگوها در دنیا (که بهم ریخته است) استفاده می‌کند. در خیلی جهات، ML پاسخی برای شکست‌های زود هنگام هوش مصنوعی سمبلیک (GOFAI) در بیرون از فضای آزمایشگاهی بود، به دلیل اینکه GOFAI قادر به پردازش پیچیدگی دنیای واقعی نبود.

الگوریتم‌های ML، با لایه‌های موازی اطلاعاتی که ارائه می‌شوند، آموزش داده می‌شوند و می‌توانند به روش‌هایی بیاموزند که نظارت نشده و نسبتا مرموز هستند! خیلی شبیه عملکرد مغز ما (از یک روش یا تابع استفاده می‌کند و آن را بر روی مجموعه‌ای از دیتا اعمال می‌کند). مانند تابعی که ایمیل‌های به درد نخور (هرزنامه) را شناسایی می‌کند؛ این تابع بر روی مجموعه‌ای از ایمیل‌ها اعمال می‌شود یا مشخص شود که کدام ایمیل به درد نخور است.

ویژگی‌های هرزنامه‌ها و غیر هرزنامه‌ها قبلا توسط انسان‌هایی که تفاوت را می‌دانند، برای الگوریتم برچسب گذاری می‌شود. از طرف دیگر، یادگیری بدون نظارت، شامل هیچ برچسب‌زنی‌ای نمی‌شود و ما نمی‌دانیم که دنبال چه فاکتورهایی هستیم! این الگوریتم در ابتدا مجموعه‌ای دیتا دریافت می‌کند و بررسی می‌کند که کدام ویژگی‌ها مرتبط هستند. برای مثال، یک الگوریتم بدون نظارت، ممکن است که به تصاویر متعددی از سگ نگاه کند و تعیین کند چه ویژگی‌هایی جوهره‌ی «سگ بودن» را به وجود می‌آورد. زمانی هم که با یک تصویر جدید روبرو می‌شود، می‌تواند تصمیم بگیرد که سگ است یا خیر.

امروزه ابزارهای علم داده خیلی کاربرپسندتر شدند و تازه‌واردان و حتی افرادی که آموزش کمی دارند، به راحتی می‌توانند وارد این زمینه شوند. این به این معنی است که هیچ وقت انجام کار با نتایج بد در این زمینه، به این آسانی نبوده است! بنابراین عواقب پروژه‌هایی بد، باید توسط کسانی که وظیفه‌ی طراحی یا اجرای آن را دارند، پیش‌بینی شود.

همانطور که کِلِهر (Kelleher) توضیح می‌دهد: «دیتا یا داده»، عنصری است که از دنیای واقعی انتزاع شده است و «اطلاعات»، داده‌هایی هستند که سازماندهی شدند تا مفید واقع شوند و «دانش» درک دقیق اطلاعاتی هست که داده‌ها به ما می‌دهند. اما با ارزش‌تر از همه، خرد است؛ که زمانی رخ می‌دهد که دانش را برای هدف خوب به کار ببریم. هدف ما این است که به خوانندگان خود کمک کنیم تا این خرد را توسعه دهند؛ که فکر می‌کنیم در قلب اخلاق علم داده قرار دارد.

بنابراین، اخلاق فقط بخشی از انجام خوب علم داده است. این یعنی، یک مشکل در دنیای واقعی، بسیار فراتر از جنبه‌های فنی آن است و البته اینکه یک سیستم چگونه قرار است زندگی افراد را تحت تاثیر قرار دهد نیز، اهمیت دارد!

۳.۱ موارد مطالعاتی

۱.۳.۱ مورد اول اخلاق تحقیق و روش علمی

مورد مطالعاتی اول، خواننده را با مفاهیمی مانند تکثیرپذیری، دقت و اعتبار آشنا می‌کند. بسیاری از این بحث‌ها بر اساس تلاش‌های اخیر در روانشناسی و همچنین علوم اجتماعی و پزشکی استوار شده است تا به واقعیتی که قسمت قابل توجهی از نتایج منتشر شده قابل تکثیر یا اعتبارسنجی نیستند، پاسخ دهند.

این مورد، سوءرفتار تحقیقاتی در آزمایشگاه غذایی کورنل Brand and Food Cornell Lab به وسیله‌ی برایان وانسینک Wansink Brian را شرح می‌دهد. مشخص شد که او برای

نتایج از چندین روش غیر علمی و البته غیر اخلاقی استفاده کرده است. روش‌هایی از جمله: picking cherry (برای علنی کردن نتایجی که مثبت بودند)، روش HAEKing (فرضیه سازی پس از مشخص شدن نتایج تجربی) و روش p-hacking (دستکاری داده‌ها برای به دست آوردن یک نتیجه آماری معنی دار)

آقایان «سینگر» و «فای تسه» تفسیری بر رفتار «وانیسنک» از دیدگاه فایده‌گرایی ارائه می‌دهند. این دو بر اهمیت راست بودن نتایج علمی که دیگران به آن تکیه می‌کنند، تأکید دارند. کسانی که این وظیفه را به عهده گرفته‌اند تا شواهد علمی و تجربی‌ای را که دیگران از آن استفاده می‌کنند، ارائه دهند، درواقع بار سنگینی را بر دوش دارند. آنها باید این کار با به بهترین نحو ممکن انجام دهند.

۲.۳.۱ زیر بخش اول

این یک مثال است.

۴.۱ بخش دوم

این یک مثال است.