فلسفهٔ علم، پژوهشگاه علوم انسانی و مطالعات فرهنگی
دوفصل‌نامهٔ علمی (مقالهٔ علمی ـ پژوهشی)، سال دهم، شمارهٔ دوم، پاییز و زمستان ۱۳۹۹، ۱۸۵–۲۰۲

یادگیری عمیق ماشینی؛ چالش های فلسفی و رهیافت‌ها

رضا نیرومند*

حمید فدیشه‌ای، الهام محمدزاده*****

چکیده

پیشرفت شگرف بشر در تولید و ذخیره‌سازی حجم انبوه داده‌ها و نیز استفاده از این مجموعه عظیم داده در ساختن ماشین استنتاج‌گر، امروزه در قالب یک فناوری پیشرفته به نام «یادگیری عمیق ماشینی» ظهور یافته است. این فناوری با الهام از اتصالات موجود در ساختار مغز جانداران طراحی شده و از شبکه‌های عصبی مصنوعی عمیق قدرت می‌گیرد. توانایی آن در استنتاج خبره‌گونه در زمینه‌های مختلف و یا تولید آثار مشابه افراد زبردست، با وجود مزایای بسیار، انسان امروز را با چالش‌های مختلفی رو در رو می‌سازد.

این نوشتار تلاش می‌کند به معرفی و تحلیل چالش‌های فلسفی و معرفتی یادگیری عمیق ماشینی و ارائه رهیافت مناسب در قبال هر یک از این چالش‌ها بپردازد. اگرچه یادگیری عمیق ماشینی چالش‌های قابل توجهی بر سر راه انسان قرار داده، با آگاهی و چاره اندیشی نسبت به آن‌ها، می‌توان در عین بهره بردن از مزایای چشمگیر این فناوری، ارزش‌های انسانی را نیز آگاهانه محافظت نمود.

کلیدواژه‌ها: یادگیری ماشین، یادگیری عمیق، شبکه‌های عصبی مصنوعی، فلسفه فناوری، چالش های فلسفی

* استادیار گروه معارف اسلامی، دانشکدهٔ علوم انسانی دانشگاه بجنورد، r.niroomand@ub.ac.ir
** استادیار گروه مهندسی کامپیوتر، دانشکدهٔ مهندسی دانشگاه بجنورد، fadishei@ub.ac.ir
*** استادیار گروه معارف اسلامی، دانشکدهٔ علوم انسانی، دانشگاه بجنورد (نویسندهٔ مسئول)، e.mohammadzadeh@ub.ac.ir

تاریخ دریافت: ۱۳۹۹/۰۳/۱۳، تاریخ پذیرش: ۱۳۹۹/۰۶/۱۲

Copyright © 2018, IHCS (Institute for Humanities and Cultural Studies). This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International, which permits others to download this work, share it with others and Adapt the material for any purpose

۱. مقدمه

در سالیان اخیر پیشرفت‌های شگرفی در فناوری در خدمت انسان قرار گرفته است که شاید چند سال قبل دور از انتظار می‌نمود. امروزه می‌توان یک دستیار انحصاری ماشینی در گوشی همراه نصب کرده با او به زبان محاوره‌ای صحبت نمود و به او فرمان داد (۱). در حین خرید از یک فروشگاه اینترنتی تعدادی کالاهای مشابه که احتمالا به آن‌ها نیاز است اختصاصا به خریدار پیشنهاد می‌شود (۲) که اتفاقا پیشنهادات کاملا مفیدی هم هستند. همین امر در هنگام گوش کردن موسیقی از یک سایت چندرسانه‌ای ممکن است به وقوع پیوندد (۳). در جایی از دنیا یک پزشک با ارسال تصاویر ماموگرافی به ابر محاسباتی(Computing Cloud) ، از مشاوره ماشین در تشخیص احتمال ابتلای فردی به سرطان سینه بهره می‌گیرد (٤) و در جایی دیگر احتمال ابتلای مردم به بیماری پارکینسون در آینده، تنها از الگوی رفتار تایپ نمودن کلمات در حین استفاده آن‌ها از گوشی همراه پیش‌بینی می‌شود (٥)! یک ماشین در گوشه‌ای دیگر در چند ثانیه یک تابلوی نقاشی خلق می‌کند که تحسین هر بیننده ای را بر می انگیزد (٦) و ماشین دیگری در حال ساخت یک قطعه موسیقی است (۷)! بدیهی است که ظهور هر فناوری نوین با خود موج جدیدی از چالش‌ها و فرصت‌ها را به همراه می‌آورد. جاذبه و هیجان نهفته در این فرصت‌ها معمولا مانع تعمق و چاره‌اندیشی و در نتیجه غلبه به موقع بر چالش‌ها می‌گردد. به همین دلیل لبه ظهور هر فناوری جدید موقعیت مهمی به لحاظ تصمیم‌گیری در جهت‌دهی آن در راستای بهره‌برداری حداکثری از فرصت‌های جدید در عین به حداقل رسانی آفت‌های ناشی از چالش‌های آن محسوب می‌شود.

مثال‌هایی که پیش‌تر ذکر گردید همگی مثال‌هایی از کاربردهای فناوری یادگیری عمیق ماشینی یا به طور خلاصه «یادگیری عمیق(Deep Learning)» هستند. این فناوری یکی از جدیدترین دستاوردهای شگرف بشر می باشد که نقش مهمی در نزدیک کردن توانمندی هوش مصنوعی(Intelligence Artificial) به هوش انسانی ایفا نموده است. هم اکنون محققان بسیاری در سراسر دنیا در دو حوزه مرتبط با این فناوری در حال تحقیق و مطالعه هستند؛ یکی حوزه پیشرفت دانش یادگیری عمیق و دیگری پهنه به کارگیری آن در کاربردهای مختلف می‌باشد.

آنچه مساله این نوشتار است چالش های فلسفی می باشد که در کنار توانمندی‌های وسیع این فناوری پدید آمده و بیشتر بر ساحت معرفتی انسانی تمرکز دارد و احتمال

خروجی های نادرست و به اشتباه انداختن آدمی را به دنبال خواهد داشت. به خطر انـداختن اختیار آدمی و عدم توجه به انتخاب گری او در اولویـت دهـی موضـوعات ماشـین و عـدم تسلط و کنترل عقلانی او در یادگیری ماشینی، انسان را از موجـودی متعقّـل و تسـخیرکننده طبیعت به مقلدی محض و مسخّر ماشین تبدیل خواهد کرد.

نویسندگان این نوشتار با توجه به نبود پژوهشی درخور این مسـاله، بـرای نخسـتین بـار گام های دقیق و بدیع پژوهشگرانه را در شناساندن دقیق چالش های فلسفی و انسـانی پـیش روی فناوری یادگیری عمیق ماشینی – که یـک فنـاوری کـاملاً جدیـد و پیشـرفته اسـت – برداشته و رهیافت لازم برای هر چالشی را بـا اتخـاذ رویکـرد فلسـفی پیشـنهاد نمـوده انـد. ساختار اصلی مقاله از سه بخش تشکیل شـده اسـت: در نخسـتین بخـش مفهـوم یـادگیری ماشین و یادگیری عمیق ارائه خواهد شـد. در بخـش دوم چـالش هـای فلسـفی پـیش روی فناوری یادگیری عمیق ماشینی مورد بحث قرار گرفته و سومین بخـش بـه ارائـه راه حـل و رهیافت‌های این چالش ها اختصاص دارد. مقاله در نهایت به جمـع‌بنـدی و نتیجـه‌گیـری از مباحث می‌پردازد.

۲. یادگیری ماشین

یادگیری ماشین(Machine Learning) به طور کلی، توانایی ماشین برای بهبود در انجام کارهـا با تکیه بر تجربه است. یعنی بر خلاف برنامه نویسـی ماشـین کـه تمـام توانـایی ماشـین بـر اساس آن چه برنامه نویس بدان حکم کرده است رقم می‌خورد، در یـادگیری ماشـین انجـام کار محوله به آن بر اساس تجربیات دریـافتی از محـیط بـه طـور پیوسـته و بـدون دخالـت برنامه‌نویس صورت می گیرد. یادگیری عمیق نوع خاصی از یادگیری ماشین است کـه ایـده کلی آن از سال‌ها قبل در ذهن دانشمندان وجود داشته ولی به دلیـل عـدم وجـود توانمنـدی فنی مورد نیاز در عملی کردن آن تا به حال، عملا تنها در سالیان اخیر به منصه ظهور رسـیده است (۸). در این نوع خاص یادگیری، با وجود منابع عظـیم داده کـه امـروزه تحـت عنـوان بیگ دیتا یا کلان‌داده‌ها(Big Data) شناخته شده اند، و با تکیه بر ماشین‌های بسیار توانمند بـه لحاظ پردازشی، امکان انجام یادگیری بسیار وسیع با استفاده از نمونه تجربه‌های بسـیار زیـاد وجود دارد. حاصل آن، ماشین‌هایی هستند کـه از توانـایی اسـتنتاج‌هـای(Reasoning) بسـیار پیچیده نزدیک به انسان خبره برخوردارند. تعریف رسمی یادگیری ماشین بـدین قـرار اسـت که وقتی کار T با توجه به یک معیار کیفیـت P بـا وجـود مجموعـه تجربیـات E بتوانـد بـه

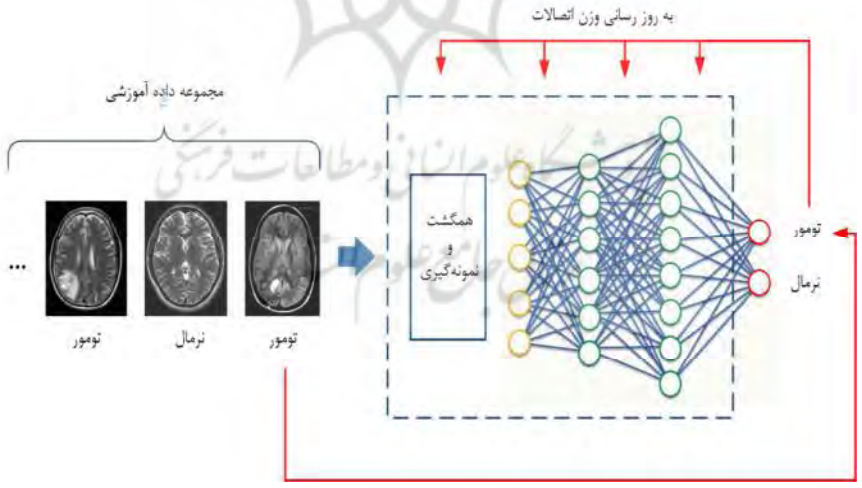
صورت بهتری نسبت به حالت بدون وجود تجربیات E اجرا شود، در این صورت یادگیری اتفاق افتاده است (۹). در بسیاری موارد ماهیت کار T در حوزه یادگیری، تشخیص می‌باشد. به عنوان مثال می‌توان به تشخیص لبخند انسان توسط یک دوربین عکاسی یا تشخیص خرابی موتور از صدای تولید شده توسط آن یا پیش‌بینی سرطان از تصاویر ماموگرافی اشاره نمود. گرچه، در عمل، T محدود به این نوع خاص نبوده و ممکن است هر نوع مسأله‌ای نظیر تولید یک اثر هنری گرافیکی یا برقراری یک دیالوگ شبه‌انسانی بین دستگاه و کاربر آن، با هدفی مشخص باشد.

رایانه‌ها از ابتدا به عنوان ماشین‌های محاسبه‌گر برای حل مسائل مختلف زندگی انسان (اقسام بی‌شمار عمل T) طراحی شده‌اند. آن چه سبب می‌شود یادگیری ماشین به عنوان یک جهش در رویکرد حل مسائل قلمداد شود، تفاوت اساسی آن با روش دیگر حل مسأله توسط ماشین؛ یعنی روش برنامه‌ریزی است. در حل مسائل به روش برنامه‌ریزی، شخص خبره در حوزه مسأله T، با همکاری متخصص برنامه‌نویسی رایانه، برنامه‌ای برای حل آن طراحی می‌کنند که با پیش‌بینی تمام ورودی‌ها و شرایط (تا حد ممکن) راه حل و گام‌های لازم برای رسیدن به پاسخ را طی می‌کند. چنین برنامه‌ای در حقیقت پیروی محض از دستورات نگاشته شده توسط این دو خبره از سوی رایانه است و رایانه تنها، مجری تک تک دستورات وارد شده می‌باشد. پس یادگیری توسط ماشین به مفهوم تعریف شده در بالا وجود ندارد و بهبود در انجام کار T تنها زمانی رخ می‌دهد که تیم برنامه‌نویسی با مشاهده تجربیات، برنامه را به لحاظ کارکردهای تعریف شده ارتقا دهند.

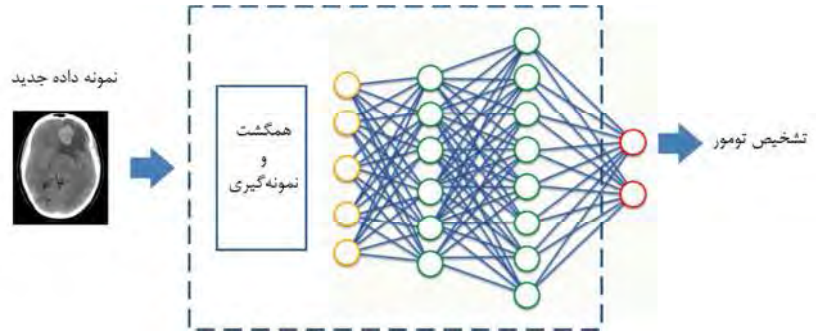
یادگیری ماشین به عنوان نقطه مقابل این رویکرد مطرح می‌شود. در این پارادایم، برنامه‌نویس تنها، یک بستر انجام تجربه و ارزیابی کیفیت انجام آن را (که به الگوریتم یادگیری شناخته می‌شود) بر روی ماشین پیاده می‌نماید. سپس این ماشین است که با دریافت تجربه‌های جدید کیفیت انجام وظیفه‌اش را بدون نیاز به برنامه‌نویسی مجدد بهبود می‌بخشد. گرچه هوش مصنوعی در هر دو رویکرد برنامه‌ریزی و یادگیری ماشین قابل دستیابی است، پیشرفت‌های دهه‌های اخیر این حوزه یقینا به میزان زیادی مدیون یادگیری ماشین است.

۳. یادگیری عمیق

یادگیری عمیق (۸) نوع خاصی از یادگیری ماشین است و بنابراین مطالبی که در خصوص یادگیری ماشین گذشت در مورد آن صادق است. وجه تمایز این نوع خاص یادگیری، استفاده از تعداد زیادی نمونه آموزشی با تکیه بر فناوری مرتبط دیگری به نام بیگ‌دیتا می‌باشد. بستر مورد استفاده در یادگیری عمیق، شبکه‌های عصبی مصنوعی عمیق یا به طور خلاصه «شبکه‌های عصبی عمیق»(Deep Neural Networks) (۸) می‌باشند که با الهام از ساختار اتصالات موجود در مغز جانداران طراحی شده‌اند. کارکرد و ساختار داخلی اتصالات بین اجزای این شبکه‌ها (نورون‌های مصنوعی) با استفاده از نمونه‌های آموزشی فراوان موجود و طی یک فرایند یادگیری استخراج می‌شود و پس از آن این شبکه‌ها قادر به انجام استنتاج برای نمونه‌های جدید داده‌ها خواهند بود. شکل ۱ مثالی از ساختار کلی و کاربرد یک شبکه عصبی عمیق را در حل مسأله پیش‌بینی تومور مغزی از روی تصاویر ام آر آی نشان می‌دهد. فرایند یادگیری ماشین شامل دو مرحله است. در مرحله اول (شکل الف) با استفاده از تعدادی نمونه آموزشی که برچسب تشخیص وجود یا عدم وجود تومور قبلا توسط متخصص به آن ها زده شده است، شبکه اصطلاحا آموزش داده می شود. در مرحله استقرار (شکل ب) شبکه آماده است که برای هر تصویر جدید تشخیص خود را ارائه نماید.



الف) آموزش شبکه با استفاده از مجموعه تصاویر برچسب خورده



ب) استقرار شبکه برای انجام تشخیص بر روی تصاویر جدید

شکل ۱) نمونه‌ای از کاربرد شبکه‌های عصبی عمیق در تشخیص تومور مغزی

٤. چالش‌های یادگیری عمیق

در این قسمت به بررسی چالش‌های فلسفی پیش روی فناوری یادگیری عمیق پرداخته شده و کوشش عقلانی برای دستیابی به راه حل و رهیافت‌های هر یک از این چالش ها در بخش بعدی ارائه شده است.

٥. احتمال استنتاج نادرست

یادگیری ماشین به طور عام، و یادگیری عمیق به طور خاص، متکی بر داده‌ها و اعداد و ارقام است. برای توسعه یک سامانه استنتاج مبتنی بر یادگیری عمیق، نیاز به تهیه انبوهی از داده‌های از قبل آماده شده وجود دارد. رویکرد کلی الگوریتم‌های یادگیری عمیق، بر کشف همبستگی‌های(Correlation) نهفته در انبوه داده‌ها میان متغیرهای ورودی و متغیر هدف پیش‌بینی استوار است. این درحالی است که همبستگی و همراهی میان دو چیز لزوما به معنای وجود رابطه علّی(Causation) میان آن دو نیست (۹). از لحاظ فلسفی، ضرورت علّی رابطه ای قطعی میان علّت و معلول است که با وجود علت، انقطاعی در ظهور معلول قابل تصور نیست. در حالی که در همبستگی، این رابطه کاملا متغیر است. به طور مثال، ممکن است یک الگوریتم یادگیری ماشین با تحلیل داده‌های یک جامعه آماری، همبستگی قابل توجهی بین میزان فقر و میزان وقوع جرم در جامعه گزارش کند، در حالی که این همبستگی لزوما به معنای این نیست که فقر عامل اصلی ارتکاب جرم باشد. ممکن است

متغیر پنهان دیگری عامل هر دو (فقر و جرم) باشد که از دید این استنتاج‌گر ماشینی پنهان مانده باشد. این متغیر پنهان ممکن است بی‌کاری یا بی‌سوادی باشد. در نتیجه، با وجود سادگی کشف همبستگی و همراهی توسط ماشین، استنتاج روابط علّی و استدلال عقلانی از توان یادگیری عمیق ماشینی خارج است.

اتکا به دقت تحلیل‌های داده‌ای ماشینی و اعتقاد به این که این نوع استنتاج می‌تواند به کلی جایگزین انسان خبره شود، با وجود نادرست بودن، حامیان خودش را داراست. این اعتقاد اصطلاحا به نام دانش عاری از تئوری(Theory-free Science) و بینش، شناخته می‌شود. جمله معروف «اعداد، خود به تنهایی سخن می‌گویند» از زبان اندرسون (۱۰) در مقاله‌ای در حمایت از این دیدگاه در زمان خود منتقدان زیادی را برآشفت. یکی از مثال‌های معروف شکست این دیدگاه در داستان واقعی «گوگل و دانش» اتفاق افتاد. زمانی مهندسان گوگل با تحلیل الگوی جستجوهای اینترنتی مراجعه‌کنندگان به وبگاه این موتور جستجو، بدون داشتن دانش پزشکی توانستند الگوی گسترش آنفلونزا در شهرهای مختلف ایالات متحده را پیش‌بینی نموده و در این زمینه به مردم اطلاع رسانی نمایند. مهم‌تر این که فرایند انجام این پیش‌بینی خیلی سریع تر از روش سنتی که توسط تیمی از متخصصان خبره با بررسی‌های میدانی انجام می‌شد به نتیجه می‌رسید. این‌جا بود که این جمله بر سر زبان‌ها افتاد که «دانش از گوگل چه می تواند بیاموزد؟» اما جالب این جاست که همین روش سریع مبتنی بر صرفا تحلیل داده‌ها بعد از چند بار پیش‌بینی به نتایج غلط و شکست انجامید در حالی که روش سنتی کماکان توسط متخصصان به نتایج صحیح در این زمینه می‌رسد.

٦. تفسیر ناپذیری خروجی

همان گونه که گذشت، مدل های یادگیری عمیق که متکی بر شبکه های عصبی عمیق می‌باشد بر اساس نمونه‌های یادگیری (تجارب از پیش مشاهده شده) شکل می‌گیرد. بر این اساس، مدل با ورود اطلاعات نمونه‌های جدید به پیش بینی لازم پرداخته و نتیجه را اعلام می‌کند. مسأله این است که نتیجه اعلام شده معمولا قابل تحلیل و تفسیر(Interpretability) به معنای رایج انسانی آن نیست (۱۱) و این امر کاربرد مدل را با چالش مواجه می‌سازد. برای روشن شدن موضوع بهتر است به مثال ذیل توجه شود. مدیران یک بانک قصد دارند برای بهبود نظام تسهیلات دست به دامان فناوری یادگیری عمیق شوند تا موفقیت یا عدم موفقیت وام گیرنده را در باز پرداخت تسهیلات اعطایی پیش بینی، و در فرایند تصمیم گیری

دخیل نمایند. فرایند مدل‌سازی یادگیری عمیق بدین نحو عمل می‌کند که سوابق نمونه‌های زیادی از وام‌گیرندگان قبلی اعم از خوش‌حسابان و بدحسابان بانک را که شامل اطلاعات مختلفی از این افراد نظیر شغل، جنسیت، درآمد، وضعیت تأهل، تعداد فرزندان و غیره می‌باشد را به عنوان تجربه (یا نمونه‌های یادگیری) استفاده می‌کند. بر اساس این نمونه‌ها، یک شبکه عصبی عمیق مختص این مسأله شکل می‌گیرد که می‌تواند در مورد متقاضیان جدید موفقیت یا عدم موفقیت در بازپرداخت را پیش‌بینی نماید. چالش مذکور در این مرحله خویش را نمایان می سازد و آن عبارت از این است که این مدل در مقام توضیح در مورد علت عدم موافقت با درخواست متقاضی خاصی، عاجز از تفسیر و پاسخ انسانی است. این در حالی است که اگر همین پیش‌بینی به روش سنتی و توسط یک تیم خبره انسانی انجام می‌شد معمولا امکان ارائه توضیح وجود داشت.

۷. اولویت دهی نامناسب و نادرست در به کارگیری

فناوری یادگیری عمیق در حوزه های مختلفی مانند سلامت، امنیت، فرهنگ و هنر، اقتصاد و سایر عرصه‌های مرتبط با زندگی انسان کاربرد دارد. برخی از این حوزه‌های کاربردی از قبیل تشخیص پیش از موعد بیماری‌ها، حائز اهمیت فراوانی هستند. از سوی دیگر، برخی از کاربردهای کم فایده نیز برای این دانش مدرن قابل تصور است. از جمله می‌توان به کاربردی تحت عنوان «جعل عمیق»(Deep Fakes) اشاره نمود که در آن چهره کاراکتر یک فیلم ویدیویی با استفاده از یادگیری عمیق با چهره شخص دیگری جایگزین می‌شود (۱۲). حال مسأله این است که با توجه به وجود درجات مختلف از میزان فایده مندی دانش یادگیری عمیق در حوزه‌های مختلف، اولویت‌دهی سرمایه گذاری فکری و مالی در این حوزه‌ها باید به چه نحو انجام پذیرد. به نظر می‌رسد آن چه امروز در جهت دهی این اولویت اثرگذار است، میزان سودآوری آن برای افراد یا شرکت‌های خاص در کسب و کارهای نوین بشری است. در حالی که چنان چه در رهیافت‌ها بدان اشاره خواهد گردید، بر اساس فلسفه فناوری به طور کلی، و یادگیری عمیق به طور خاص، محور این اولویت‌گذاری را باید کرامت انسانی(Human Dignity) تلقی نمود.

۸. مداخله در اراده آزاد انسان (اختیار)

بسیاری از فیلسوفان اراده آزاد(Free Will) آدمی را وجه تمایز انسان از حیوان دانسته تا آنجا انسان را به حیوان مختار تعریف نموده‌اند. از این رو اختیار به عنوان یکی از ذاتیات انسانی در اخلاق، حقوق و سایر شئونات زندگی او از اهمیت فوق‌العاده‌ای برخوردار است. هر چیزی که با این اصل انسانی در تعارض قرار گیرد می‌تواند به عنوان تهدیدی بر انسانیت انسان قلمداد شده و لذا مردود بودن آن از نگاه جامعه انسانی قطعی است.

فناوری یادگیری عمیق این توانایی را به برخی افراد و گروه‌ها می دهد که مخاطب و جامعه را به سمت و سوی خاصی سوق دهد. به طور نمونه تبلیغات نمایش‌یافته در حاشیه یک وب سایت خبری می‌تواند به گونه‌ای آرایش یابد که شخص مخاطب را به سمت رأی دادن به نامزد انتخاباتی خاصی سوق دهد. یا دنباله ویدئوهای پیشنهادی توسط ماشین در یک سایت چند رسانه‌ای می‌تواند به گونه‌ای طرح‌ریزی شده باشد تا به تدریج فرد را به گرایش جنسی خاصی متمایل کند. البته این مداخله در انتخابگری ممکن است لزوما در موارد منفی و غیر اخلاقی رخ ندهد. مثلا ممکن است هدف از دنباله ویدئوها در مثال قبل علاقمندسازی شخص به محیط زیست باشد. اما به هر صورت مداخله در اختیار انسانی در این حالت نیز به وقوع پیوسته است. این امر از گذشته تحت عنوان «گرایاندن نامحسوس»(Nudge) یا «پدرخواندگی نرم»(Soft Paternalism) شناخته شده است و از مصادیق آن می‌توان به گذاشتن سالاد در کنار غذا برای ارتقای سلامتی اشاره نمود. اگرچه در دوران یادگیری عمیق با نوع جدید و بسیار قدرتمندتری از این امر مواجه هستیم که برای آن واژه «ابَرگرایاندن»(Hypernudge) پیشنهاد شده است (۱۳). این قدرت آنچنان بزرگ است که اکنون این سؤال برای بشر متفکر به وجود آمده است که آیا دموکراسی از مهلکه هوش مصنوعی و کلان‌داده‌ها جان سالم به در خواهد برد یا نه (۱٤)!

از دیگر سو، آن چه اختیار آدمی را تحت‌الشعاع قرار می‌دهد، فراگیری دستیارهای دیجیتال است. این دستیارها، برنامه‌های کاربردی قابل اجرا بر روی گوشی همراه افراد است که قادرند در امور روزمره زندگی، انتخاب و پیشنهاد بهینه در تصمیم گیری هایی نظیر رژیم غذایی، مسیر حرکت، برنامه تفریح و نظایر آن را به شخص ارائه نمایند. غلبه این دستیارها از سویی، و پیشرفت تصمیمات آن‌ها در سایه یادگیری ماشین از سوی دیگر، عملا افراد را به عادت به حضور و اتکا به وجود آن در زندگی سوق داده و منجر به کاهش تجربه انتخاب در افراد می‌شود (۱۵).

۹. اثرگذاری نامطلوب بر توسعه و رشد فکری و اجتماعی

ورود ماشین‌های هوشمند در ابعاد مختلف زندگی متضمن نوعی کنترل روز افزون بر انسان می‌باشد. این روند با بلوغ تکنولوژی به سمتی پیش می‌رود که از ارتکاب حتی کوچک‌ترین اشتباهات نیز جلوگیری شود. به عنوان نمونه، امروزه یک راننده در یک کشور پیشرفته ممکن است به خاطر سرعتی معادل یک کیلومتر بر ساعت بیشتر از حد مجاز جریمه گردد، یا شخصی توسط دستیار دیجیتال خود از خوردن یک وعده غذایی جدید در موقعیتی خاص به دلیل عدم تطابق با رژیم مناسبش منع گردد. این تصمیمات پشتوانه کاملا علمی و فنی داشته، نتیجه علم بشری بوده و برای رسیدن به نتایج ایده ال طراحی شده است. در عین حال این میزان دقت علمی و فنی وقتی که با نوعی تلاش برای به صفر رساندن خطاهای بشر و تصمیمات نادرست همراه است، عملا باعث حرکت به سمت جامعه همگون و یکسان سازی جامعه و افراد خواهد شد. این نوع همگونی علاوه بر این که مانعی بر سر راه توسعه فکری افراد و خلاقیت آن‌هاست، عاملی برای از بین رفتن خرد جمعی و توسعه اجتماعی بشر خواهد بود. چرا که خرد جمعی حاصل دگر اندیشی و برآیند تفاوت در دیدگاه‌ها و افکار جامعه است (۱۵).

۱۰. نقض حریم خصوصی

همان گونه که در بالا گذشت، یادگیری معادل بهبود از طریق تجربه تعریف می شود. در یادگیری ماشین تجربه را می‌توان معادل داده و در یادگیری عمیق معادل کلان‌داده دانست؛ به عنوان مثال برای پیش‌بینی وقوع جرم در یک محدوه جغرافیایی با استفاده از این فناوری نیاز به داشتن مجموعه‌ای از داده‌های مرتبط با تاریخچه جرم و جنایت در آن منطقه است. چنین پایگاهی می‌تواند حاوی اطلاعات خصوصی شهروندان باشد و دسترسی همگانی به آن، نقض حریم خصوصی(Privacy) افراد را در پی داشته و زمینه سوء استفاده را فراهم سازد.

آنچه این موضوع را پررنگ تر می‌سازد این حقیقت است که الگوریتم‌های یادگیری عمیق عمدتاً پیچیدگی محاسباتی بالایی دارند و از آنجا که رایانه‌های معمولی قادر به پردازش آنها نمی باشند اجرای آنها به سرورهای قدرتمند که بر فضای ابر محاسباتی توزیع شده اند محول می‌گردد. این امر سبب پراکندگی و نقل و انتقالات متعدد داده ها از مرحله

جمع آوری تا تحلیل در فضای اینترنت خواهد شد که احتمال سوء استفاده و دستیابی نفوذ گران را به داده‌های مذکور بیشتر می‌سازد.

توجه به این نکته لازم است که مسأله حریم خصوصی معمولا نه به صورت صفر و صد (مخفی‌کاری کامل اطلاعات یا آشکارسازی کامل اطلاعات) بلکه چیزی بین این دو حالت تعریف می‌شود. تن دادن به نگاه صفر و صدی سبب می‌شود که مجبور به پذیرش این موضوع شویم که باید از بین مزایای فناوری یادگیری عمیق و حفظ حریم خصوصی انسان‌ها یکی را انتخاب کنیم. حریم خصوصی در حقیقت وجود سطوح تعریف شده در دسترسی به جریان اطلاعات است (۱۷).

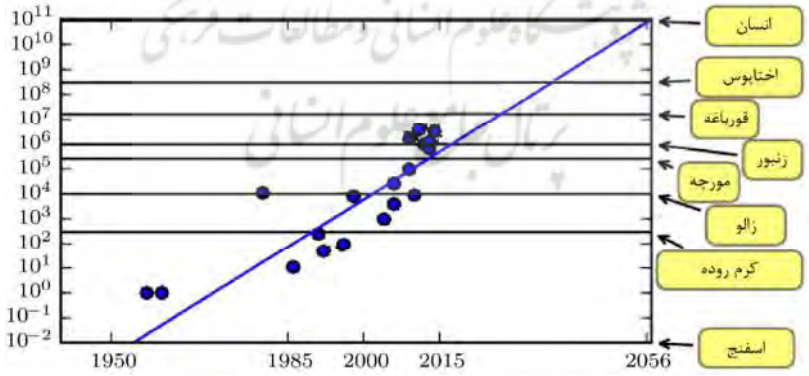
۱۱. رهیافت‌های پیش رو

اگر چه این نوشتار در شمار نخستین گام های جدی جامعه علمی ما در مواجهه با چالش های فناوری پیش‌رونده یادگیری عمیق ماشینی است، اما به نظر می رسد باید در جهت حل چالش‌های مذکور تلاش عقلی، فلسفی و علمی صورت گیرد. از این رو نویسندگان بر آن شدند تا حد امکان در جهت حل آن، رهیافت‌های زیر را پیشنهاد کنند. باید توجه داشت که گرچه ممکن است نتوان رهیافتی اساسی و راه حلی نهایی برای این چالش‌ها یافت، اما گام برداشتن در این راستا نیز بی‌فایده نخواهد بود.

۱۲. احتمال استنتاج نادرست

همان‌گونه که گذشت، چالش موجود در یادگیری ماشینی عمیق در عدم توجه به تفاوت ماهوی همبستگی و علیت است که نتیجه آن احتمال خروجی نادرست را رقم می‌زند. در مقابل چالش مذکور ممکن است این پرسش به ذهن خطور کند که آیا با توجه به روند پیشرفت تکنولوژی یادگیری ماشین، دستیابی به توانایی استنتاج علّی توسط ماشین، آنچنان که از یک خبره انسانی توقع می‌رود، امری قابل دستیابی در آینده نزدیک نیست؟ به لحاظ فنی، رسیدن یادگیری ماشین به توان استنتاج علّی، نیازمند دو عامل است: وجود ماشین یادگیر توانمند در پردازش داده‌های فراوان، و وجود داده دقیق، مرتبط، و کافی برای استفاده توسط این ماشین.

تاریخ پیشرفت فناوری ، نشانگر پیشرفت ماشین‌های یادگیر در طول زمان می‌باشد. شکل ۲ این پیشرفت را به صورت روند افزایش پیچیدگی شبکه‌های عصبی مصنوعی ساخته شده توسط انسان – که زیرساخت اصلی در یادگیری ماشینی عمیق هستند – در طول زمان نشان می‌دهد (۸). این تصویر ضمنا مقایسه‌ای از این تعداد بین شبکه‌های عصبی مصنوعی و مغز برخی جانوران مختلف را به تصویر کشانده است. تصویر نشان می‌دهد که در صورت ادامه روند فعلی، تخمین زمان دستیابی به شبکه عصبی مصنوعی با تعداد نورون برابر با مغز انسان سال ۲۰۵۶ خواهد بود. البته این بدان معنا نیست که قدرت استنتاج چنین شبکه‌ای لزوما برابر با مغز یک انسان خواهد بود اما به هر حال در پاسخ به پرسش مذکور، عامل پیشرفت ماشین قابل دستیابی بوده و امری دور از انتظار نیست. اما عامل دوم، یعنی وجود داده دقیق و کافی که مرتبط با مسأله باشد، نیازمند حضور انسان خبره در فرایند یادگیری و نظارت بر آن هست. این حضور هم اکنون در قالب پژوهش‌هایی در زمینه یادگیری ماشینی و در حوزه‌هایی نظیر تفسیرپذیر سازی مدل‌های یادگیری و ایجاد مدل‌های دوگانه با امکان داخل نمودن تجربیات خبره در مدل به طور صریح، توسط محققان علوم و مهندسی رایانه در جریان است. بنابراین بعید به نظر نمی‌رسد که با توجه به فراز و فرودی که در پیشرفت تکنولوژی به دست بشر شاهد بوده‌ایم، روزی فرا رسد که تمام آن چه نیازمند انتقال قدرت استنتاج خبره انسانی به ماشین است فراهم گردد. در عین حال، اگر چه این انتقال قدرت استنباط به ماشین از لحاظ فلسفی ممکن است، اما نباید این نکته اساسی را دور از نظر داشت که این جریان انتقال امری پیوسته است بدین معنا که همیشه در کنار پیشرفت ماشین، پیشرفت و پرورش خبره انسانی نیز امری ضروری است.



شکل ۲) پیشرفت شبکه‌های عصبی مصنوعی ساخته بشر در طول زمان به لحاظ تعداد نورون تعبیه شده در آن‌ها. هر نقطه آبی‌رنگ مربوط به یکی از شبکه‌های عصبی مصنوعی ساخته محققان است (۸)

۱۳. تفسیر ناپذیری خروجی

اگرچه اغلب مدل‌های یادگیری عمیق در حال حاضر فاقد توان ارائه تفسیر بـر اسـتنتاج‌هـای خویش هستند، نیاز به رفع این نقص نمایان شده است و محققانی در حال کاوش و تحقیـق با هدف تفسیرپذیر نمودن مدل‌های یادگیری عمیق هستند. به عنـوان مثـال در مـدلی کـه بـا هدف تشخیص بیماری سرطان از تصاویر رادیـوگرافی، ممکـن اسـت یـک مـدل یـادگیری عمیق به گونه ای طراحی شود که قادر به تجزیه عوامل اثرگذار تصویر (به عنوان مثال محـل توده های مشکوک به سرطان) و متمایز نمودن آن ها و اثرات آن‌ها بر استنتاج باشد.

به علاوه، نکته حائز اهمیت این است که یـک تـلازم پـر رنـگ میـان تفسـیرپذیر بـودن خروجی ماشین و قدرت استنتاج و در نتیجه حضـور خبـره در فراینـد یـادگیری و اسـتنتاج وجود دارد. تفسیرپذیر بودن به معنای انسانی که مد نظر ماست زمـانی محقـق مـی‌شـود کـه مفسر از قدرت استنتاج برخوردار باشد و از آنجایی که در بحث قبلی قدرت اسـتنتاج مـورد چالش قرار گرفته و نیاز به حضور خبره نمایان گردیـد، بنـابراین بـا رهیافـت ذکـر شـده در مسأله قدرت استنتاج، می‌توان چالش تفسیرناپذیری خروجی را نیز تا حدی حل نمود.

۱۴. اولویت دهی نامناسب و نادرست در به کارگیری

همان گونه که پیش‌تر ذکر شد، یادگیری عمیق در زمینه‌های مختلفی کاربرد دارد که این امـر چالش اولویت‌بنـدی در هـر یـک از زمینـه‌هـای ممکـن را رقـم مـی‌زنـد. بـه طـور کلـی، اولویت‌بندی مؤثر باید در سطح سیاست‌گذاری کلان جوامع انجـام شـود. امـا ایـن پرسـش مطرح می‌شود که معیار مناسب این اولویت بندی چیست؟

به باور نگارندگان، کرامت انسانی می تواند به عنوان محور اصلی این اولویت بنـدی قـرار گیرد. به عبارت دیگر، در توسعه زمینه‌های مختلف ایـن فنـاوری، هـر مـورد کـه همسـویی بیشتری با کرامت انسانی دارد در اولویت بالاتری بـرای پـرداختن و رشـد نسـبت بـه سـایر زمینه‌ها قرار گیرد. در توسعه فناوری یادگیری عمیق به صورت موردی بر این اسـاس امکـان تدوین سنجه‌های ثانوی از این محـور اصـلی وجـود خواهـد داشـت. بـه عنـوان مثـال، در بودجه‌ریزی برای حمایت از تحقیقات دانشگاهی، می‌توان ابتدا به دنبال این مسأله رفـت کـه دستاورد این تحقیق در نجات جان چند انسان یا رفع فقر چند نفر اثرگذار خواهد بود. یـا در

انتخاب بین دو تحقیق موردی که یکی به تشخیص زودهنگام سرطان و دیگری به جعل عمیق چهره افراد در ویدئوها می‌پردازد، مورد اول در اولویت اقدام قرار گیرد.

۱۵. مداخله در اراده آزاد انسان (اختیار)

همان گونه که گذشت اختیار و اراده آزاد انسانی وجه ممیز انسان می باشد که در نهاد و ذات او به ودیعه گذاشته شده است. فرایند عملی اختیار انسان آن چنان نیست که بتواند به راحتی دستخوش نابودی گشته و یا حتی بتواند مورد خدشه و مداخله قرار گیرد. تاثیر گذاری و سمت سوی دادن عادت و سلیقه و حتی اندیشه افراد در تصمیم سازی اختیار آدمی امری طبیعی است و طبیعتا انسان با توجه به عوامل متعدد و مختلف نهایتا تصمیم گیری می‌کند. باید توجه داشت که خارج از فناوری یادگیری عمیق، انسان تحت تاثیر ژنتیک، نیاز های فیزیولوژیک، عقاید بنیادین و دینی، جامعه و محیط تربیتی و بسیاری عوامل دیگر به انتخاب دست می‌زند. با وجود عوامل بی‌شماری که در اختیار آدمی تأثیر گذارند، آخرین عامل اثرگذار در انتخاب و انجام یک فعل اختیار خود فرد است.

برای مواجهه با این چالش، دو راهکار اساسی پیشنهاد می گردد. یکی از این دو راهکار، تغییر رویه از «گرایاندن نامحسوس» به «گرایاندن محسوس» است (۱۵). نمود این پیشنهاد در مثالی که قبلا ذکر شده بود، این است که مثلا ذیل یک تصویر ویدئویی ذکر شود که آگاه باشید دیدن این ویدئو ممکن است شما را تبدیل به یک فرد علاقمند به محیط زیست نماید. راهکار دوم، توسل به مواجه سازی کاربران با تضادها و استفاده از اثر شگرف این امر بر رشد انسان است. اگر ماشین ها جهات مختلفی به خصوص دو سوی مقابل را پیش روی ادمی بگذارند و پیشنهاد کنند، در واقع این خود مصداقی برای ظهور کامل اختیار ادمی در این فضا می باشد. به عنوان مثال، در زمانی که شخصی در یک سامانه محتوای ویدئویی با دنباله‌ای از ویدئوها با مضمون‌های محرک جنسی مواجه شده، در این میان به طور ناگهانی یک ویدئو در مورد اهمیت کنترل این نیرو و جهت دهی آن در راستای صحیح گنجانده شود. این تضادها به مثابه خروجی‌های تعبیه شده در یک بزرگ‌راه عمل خواهند کرد تا افراد از جبر راه سپردن در یک مسیر خاص رهایی یابند.

۱۶. اثرگذاری نامطلوب بر توسعه و رشد فکری و اجتماعی

پیش فرض این چالش مبتنی بر وجود جبر آدمی در دنیای فناوری یادگیری عمیق است. این پیش فرض بر اساس رهیافتی که در باب اراده آزاد آدمی گذشت مبنای صحیح و دقیق نداشته و بیش از انکه علمی و منطقی باشد بر اساس نوعی بدبینی به فناوری های جدید طرح گشته است. در عین حال همان‌گونه که گذشت، جامعه‌ای که در چشم‌اندازش ابَرهوش مصنوعی و یادگیری عمیق را پیش رو قرار داده است، جامعه‌ای است که افراد در آن (اغلب به شیوه‌های نامحسوس) تحت نظارت بوده و از ارتکاب کوچکترین خطاها نیز جلوگیری می‌شود. در چنین جامعه یکدستی که هم اکنون نیز علائمی از آن در حال ظهور و مشاهده است، همه شهروندان کمابیش شبیه یکدیگر هستند و خرد جمعی که نتیجه تفاوت آراء است رنگ می‌بازد و توسعه فکری و اجتماعی بشر دچار کندی می‌گردد.

برای مقابله با این چالش، اولین قدم به رسمیت شناختن و تذکر دادن وجود اراده آزاد به شهروندان جامعه است. کاستن از میزان اتکای افراد به ماشین و استنتاج‌های آن نیز چاره مهم دیگری در این مسیر است. همان‌طور که برای حفاظت از طبیعت روز بدون خودرو پیشنهاد می گردد، ممکن است در دنیای مدرن هوش مصنوعی، روزی یا ساعتی بدون هوش ماشینی حداقل برای جزئی از ابعاد نفوذ این فناوری در زندگی افراد پیشنهاد گردد تا انسان بتواند لحظاتی خاص و در مواردی خاص به خود و قدرت تصمیم و تجربه خود بازگردد. نباید فراموش کرد که روزهای نه چندان دوری در گذشته نزدیک، شهروندان شماره تماس نزدیکان خود را در ذهن داشتند، مسیریابی حرکت خود در خیابان‌های شهر را در ذهن خود رقم می‌زدند و در انتخاب قطعه موسیقی مورد علاقه خود متکی به نظر ماشین نبودند. پس این امر غیر ممکن نیست. البته ممکن است این امر مسبب ارتکاب اشتباهاتی یا تجربیات ناخوشایندی گردد، اما باید توجه داشت که بسیاری از دستاوردهای مهم در تاریخ پیشرفت بشر از ابتدا حاصل یک اشتباه بوده است.

۱۷. نقض حریم خصوصی

چنان که در مقدمه این نوشتار ذکر شد، فرایند یادگیری عمیق نیاز به نمونه‌های زیادی از داده‌ها برای یادگیری دارد. به همین دلیل این فناوری ارتباط تنگاتنگی با فناوری بیگ‌دیتا دارد. جمع‌آوری داده از فعالیت‌های روزمره انسان‌ها از حضور آن‌ها در شبکه‌های اجتماعی

گرفته تا رفت و آمد و سوابق پزشکی آن‌ها گرچه می‌تواند منجر به دستاوردهای مفیدی با کمک یادگیری عمیق شود، مخاطره نقض حریم خصوصی آن ها را نیز به دنبال دارد. فهرستی مختصر از رهیافت هایی در مواجهه با نقض حریم خصوصی به شرح زیر پیشنهاد می‌گردد (۱۷):

- شفافیت: آگاهی از جزئیات اطلاعات ذخیره شده کاربران و امکان اختیار آن چه ذخیره می‌شود و آن‌چه مجاز به ذخیره شدن نیست باید جزو حقوق کاربران تلقی شود.
- سطوح دسترسی: دسترسی به اطلاعات کاربران باید در سطوح مختلفی محدود شده به نحوی که در هر سطح به میزان کافی و نه بیشتر از آن دسترسی وجود داشته باشد.
- توسعه قوانین: قوانین پویا و کارآمد برای حمایت از حریم خصوصی افراد باید تدوین گردد.
- توسعه فنی: محققان باید به تکاپو برای امن ساختن محیط های کاربری اهتمام ورزند. برخی روش‌های پردازش داده ها که با داده‌های بی‌نام شده، مبهم شده، یا تکه تکه و توزیع‌شده به جای داده‌های اصلی کار می‌کنند توسعه یابند.

۱۸. نتیجه‌گیری

همان طور که به تفصیل گذشت فناوری یادگیری عمیق ماشینی مانند هر پدیده نوظهوری که محصول پیشرفت علم و فناوری است طبیعتاً با مسائل و چالش های فلسفی، اخلاقی و اجتماعی روبرو است. کانون و دغدغه اساسی چالش های مذکور که عمدتاً رنگ و بوی فلسفی، منطقی و معرفتی دارند ساحت شناختی و در نتیجه آن، ساحت رفتاری انسان را از جهت درست و نادرست مورد هدف قرار داده است. علم و معرفت و به طریق اولی فناوری از نگاه انسان شناختی زمانی ارزشمند محسوب می شود که آدمی را از جهل به علم و از بیهودگی به فایده مندی برساند. از این رو، پژوهش حاضر تلاش نموده به دو مساله اساسی پرداخته و پاسخ دهد: چالش‌های فلسفی این فناوری پیشرفته چیست؟ و چه راهکارهایی در رویارویی با این چالش‌ها وجود دارد؟ در این مقاله چالش‌های موجود با دیدگاهی فلسفی مورد تبیین قرار گرفته و در ادامه راهکارهایی برای آماده شدن در مقابله با این چالش‌ها ارائه گردیده است. به عنوان نتیجه حاصل از این پژوهش باید گفت با توجه

به راه حل هایی که گذشت می توان چالش های نهفته در این فناوری را همواره بـا کوشـش عقلانی و علمی حل کرده و در نتیجه باعث شـد از سـوق دادن آدمـی بـه خطـا جلـوگیری نموده تا با نتایج و پیامدهای مثبت آن باعث بهبود زندگی بشری شود.

با توجه به این که مسیر چالش ها و رهیافت ها امـری مسـتمر و مسـیری تمـام نشـدنی است، مقابله با این چالش‌ها نیازمنـد کوشـش هـر دو طیـف پژوهشـگران در حـوزه نظریـه پردازی فلسفی و علوم فنی و کاربردی رایانه برای تبیین این چالش ها و کوشش برای یـافتن راه حل های نظری و عملی و ایجاد بستر مناسب برای مقابله با آن‌هاست. بنابراین بـه جـای دو راهبرد آسان و غیرمعقـول بـدبینی، محدودسـازی و تعطیلـی فنـاوری از سـویی و انکـار چالش های انسانی و معرفتی فناوری ها از دیگر سو، باید دائما به تعقل، تفلسف و کوشـش علمی و فنی در جهت شناسایی و حل مسائل و چالش های مذکور پرداخت.

کتاب‌نامه

Sarikaya, Ruhi (2017). "The technology behind personal digital assistants: An overview of the system architecture and key components." IEEE Signal Processing Magazine 34.1: 67-81.

Jannach, Dietmar, and Malte Ludewig (2017). "When recurrent neural networks meet the neighborhood for session-based recommendation." Proceedings of the Eleventh ACM Conference on Recommender Systems. ACM.

Van den Oord, Aaron, Sander Dieleman, and Benjamin Schrauwen (2013). "Deep content-based music recommendation." Advances in neural information processing systems.

Sun, Wenqing (2016). et al. "A preliminary study on breast cancer risk analysis using deep neural network." International Workshop on Breast Imaging. Springer, Cham.

Adams, Warwick R (2017). "High-accuracy detection of early Parkinson's Disease using multiple characteristics of finger movement while typing." PloS one 12.11: e0188226.

Huang, Xun, and Serge Belongie (2017). "Arbitrary style transfer in real-time with adaptive instance normalization." Proceedings of the IEEE International Conference on Computer Vision.

Agarwala, Nipun, Yuki Inoue, and Axel Sly (2017). "Music composition using recurrent neural networks." CS 224n: Natural Language Processing with Deep Learning, Spring.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville (2016). Deep learning. MIT press.

Mitchell (1997). Machine learning, McGraw Hill.

Wright, Sewall (1921). "Correlation and causation." Journal of agricultural research 20.7: 557-585.

Anderson C. (2008). The end of theory: the data deluge makes the scientific method obsolete. Wired Magazine; 16(7).

Zhang, Quan-shi, and Song-Chun Zhu (2018). "Visual interpretability for deep learning: a survey." Frontiers of Information Technology & Electronic Engineering 19.1 :27-39.

Harris, Douglas (2018). "Deepfakes: False Pornography Is Here and the Law Cannot Protect You." Duke L. & Tech. Rev. 17–99.

Yeung, K (2017). 'Hypernudge': Big Data as a mode of regulation by design. Information, Communication & Society; 20(1): 118-136.

Helbing, D. et al (2019). Will democracy survive big data and artificial intelligence? Towards Digital Enlightenment. Springer, Cham; 73-98.

Helbing, D (2019). Societal, economic, ethical and legal challenges of the digital revolution: From big data to deep learning, artificial intelligence, and manipulative technologies. Towards Digital Enlightenment. Springer, Cham; 47-72.

Fadishei. H, Niroomand. R, Mohammadzadeh. E (2017). Ethical Challenges of Big Data

Vol. 11, No. 4: 25-34 (In Persian).

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی