

۱- دکتر امیرحسین بحرالعلومیان؛
دکتری تخصصی مهندسی پزشکی
۲- مهندس منیره توکلی (کارشناس ارشد
مهندسی پزشکی)



کاربرد هوش مصنوعی در تشخیص های آزمایشگاهی - بخش دهم

چالش ها

علیرغم پیشرفت های زیادی که در زمینه هوش مصنوعی اتفاق افتاده و با استفاده از داده های عظیم، قدرت محاسباتی بالا و شبکه های عصبی پیشرفته، کیفیت مطالعات مربوط به آنالیزهای روتین خون و ارتباط آن ها با تشخیص و پیش آگهی بیماری ها بهبود پیدا کرده، اما هنوز استفاده از این روش ها در معالجات بالینی یک چالش بزرگ است. مطالعات و بیماری هایی که بررسی کردیم، تأیید می کنند که در پیاده سازی فناوری های مبتنی بر هوش مصنوعی در محیط بالینی تأخیر وجود دارد.

وجود اطلاعات آماری قابل توجهی در این مرور است که به دنبال بررسی ارتباطات مهم بین ترکیبات خونی و بیماری ها است و از فرصت های موجود در مراکز درمانی بهره می برد. همان طور که اشاره شد، این پنل های تحلیلی شامل شمارش کامل سلول های خونی، آزمایش های متابولیک و لیپیدی است که در حال حاضر با استفاده از تکنیک های استاندارد و بسیار پایدار ارزیابی می شود و به خطاهای سیستماتیک یا بایاس حساس نیست. مقدار زیادی از اطلاعات بالینی که تاکنون مورد بررسی قرار نگرفته، آنقدر زیاد است که یک پزشک به تنهایی نمی تواند در طول یک ویزیت معمولی یا حتی در شرایط اورژانسی، همه آن ها را به طور کامل درک کند (به خصوص در مواردی که نیاز به پیگیری طولانی مدت باشد). اما این اطلاعات می تواند با استفاده از فناوری های پیشرفته پردازش شود، الگوهای مشخصی در آن ها شناسایی گردد، از نظر آماری مورد تجزیه و تحلیل قرار بگیرد و در صورت لزوم به پزشک هشدار داده شود. تصمیمات بالینی فعلی بر اساس سیستم های مبتنی بر قواعد گرفته می شود، به این معنی که آستانه های مشخصی وجود دارد که مطابق با دستورالعمل های جدید به روز

می شود. دو دلیل اصلی که باعث مقاومت در برابر استفاده از راه حل های مبتنی بر یادگیری ماشین می شود:

۱- نیاز به استفاده از برنامه های کاربردی خارجی: این برنامه ها نیاز به وارد کردن دستی داده ها دارند و باعث اتلاف وقت می شوند.

۲- عدم قابلیت تفسیر الگوریتم های یادگیری ماشین: به ویژه الگوریتم های مربوط به یادگیری عمیق که به عنوان 'جعبه سیاه' شناخته می شوند.

به عبارت دیگر، پزشکان ترجیح می دهند از سیستم های آشنا و قابل تفسیر استفاده کنند، به جای اینکه مجبور به استفاده از برنامه های جدید و پیچیده ای باشند که داده ها را به صورت دستی وارد کنند و نتایج آن ها را هم نمی توانند به خوبی درک کنند.

مطالعه ای که به تازگی توسط هنری و همکاران انجام شد، به بررسی یک سیستم هشدار زود هنگام مبتنی بر هوش مصنوعی برای تشخیص سپسیس (TREWS) پرداخت. این محققان متوجه شدند که مدل محاسباتی این سیستم قابلیت تفسیر خوبی ندارد، اما این موضوع مانع مهمی نبود. به ویژه، پس از اینکه پزشکان تجربه کار با این سیستم را در بیماران مختلف داشتند و با همکاران و اعضای تیم تحقیق تعامل کردند، نگرانی های اولیه آن ها کاهش یافت.

از سوی دیگر، تشخیص های نظری «رقابتی» ممکن است توسط برخی پزشکان برای استقلال عمل آن ها برای تشخیص را تحت تأثیر قرار دهد. به همین دلیل آن ها در پذیرش این راه حل ها تردید دارند، زیرا ممکن است فرآیند تصمیم گیری آن ها را تغییر دهد و خطر این وجود داشته باشد که تنها براساس توصیه های مدل، که ممکن است کاملاً دقیق نباشند، عمل

کنند. در مورد تحلیل‌های روتین خون، منابع داده‌ای (مانند تجهیزات، شیوع بیماری، و ویژگی‌های دموگرافیک بیماران) با مقادیر مرجع متفاوت نیز باید مورد ارزیابی و بحث قرار گیرد. این مطالعه باید به‌طور واضح اطلاعاتی در مورد نوع منبع داده (مانند گروه‌های هم‌گروه، کارآزمایی‌های کنترل‌شده تصادفی یا سایر انواع) ارائه دهد. همچنین کیفیت منبع داده (شامل نمایندگی، سوگیری، ویژگی‌ها و نتایج با زمان دقیق اندازه‌گیری و دارو یا درمان مرتبط) و همچنین مقدار منبع داده را نیز باید گزارش کند. ویژگی‌های منحصربه‌فرد بیوشیمیایی افراد در شرایط پایه به دلایل مختلفی متغیر است و در بیشتر موارد این تغییرات ارتباطی با وضعیت بالینی آن‌ها ندارد. مطالعات مجدد می‌توانند در نهایت مشخص کند که آیا پیش‌بینی‌های مدل هوش مصنوعی به ویژگی‌هایی مرتبط با مشکل مورد نظر ارتباط دارند یا اینکه فقط متغیرهای خارجی مانند نوز حسگر، دماهای محیطی، دستکاری کاربر و موارد مشابه را تحت تأثیر قرار می‌دهند. (استگمان و همکاران، ۲۰۲۰)

با توجه به ارزیابی برنامه‌های پزشکی بررسی‌شده، تنها تعداد کمی در مراکز خارجی انجام شده‌اند و بیشتر آن‌ها با داده‌های گذشته‌نگر (رتروسپکتیو) صورت گرفته است. بنابراین، کار با داده‌هایی که شرایط مشابه محیط‌های بالینی سنتی را داشته باشند، امری حیاتی است. به ویژه در مواردی مانند رابط کاربری (برای متخصصان بهداشتی یا بیماران) و یکپارچه‌سازی فناوری در جریان کاری بالینی (شرایط فیزیکی مانند نور، دما، رطوبت و غیره).

اخیراً یک ارزیابی برای پیش‌بینی عملکرد یک سیستم یادگیری عمیق برای تشخیص رتینوپاتی دیابتی انجام ده است که نشان داد تعداد زیادی از تصاویر شبکه به دلیل تاری یا تیرگی، به عنوان غیرقابل ارزیابی شناخته شده‌اند. این مشکل ناشی از نور ضعیف محیطی در طول فرایند اندازه‌گیری بوده است.

در مورد مدل‌های پیش‌بینی COVID-19، یک مطالعه نشان داد که توزیع داده‌های زیرین (Underlying)، که به عنوان شیفت دامنه شناخته می‌شود، تأثیر قابل توجهی بر عملکرد و قابلیت اعتماد پیش‌بینی‌شده دارد و منجر به شکست مدل در کاربردهای بالینی می‌شود. شیفت دامنه، که می‌تواند ناشی از تغییرات در شیوع بیماری، تنظیمات پروتکل‌های آزمایش RT-PCR یا جهش‌های ویروسی باشد، نشان می‌دهد که مدل‌های یادگیری ماشین ممکن است به مرور زمان قابلیت اطمینان و عملکرد خود را از دست بدهد. این موضوع بر اهمیت نظارت و به‌روزرسانی مداوم تأکید می‌کند. (رولند و همکاران، ۲۰۲۲) این مثال‌ها تأکید می‌کنند که آموزش باید

شرایط اصلی را در نظر بگیرد تا ضرایب واقعی برای حل مشکل مورد نظر ایجاد شود.

این چالش‌ها باید با پیاده‌سازی مداوم pipeline یادگیری ماشین توصیف‌شده، بازنگری و بهبود یابند تا یادگیری فدرال (آموزش در چندین مؤسسه) و استقرار در ETL (یعنی استخراج، تبدیل و بارگذاری) توسعه یابد و داده‌ها «سالم» باقی بمانند.

چشم انداز آینده

ابزار تشخیصی که بر اساس داده‌های طولی (داده‌های جمع‌آوری‌شده در طول زمان) و استانداردهای پایدار (معیارهای مرجع) طراحی شده، قرار است به صورت رایگان و در زمان واقعی به کار گرفته شود.

هدف از استفاده از این ابزار، ارائه یک تصویر کلی و جامع از وضعیت سلامت هر فرد به پزشک است. به عبارت دیگر، این ابزار با استفاده از داده‌های معتبر و به‌روز، احتمال ابتلا به بیماری‌های مختلف را برای هر فرد محاسبه می‌کند و نتایج را به صورت گزارشی جامع در اختیار پزشک قرار می‌دهد. این امر به پزشک کمک می‌کند تا تصمیم‌گیری‌های بالینی دقیق‌تری داشته باشد.

هوش مصنوعی می‌تواند نقش کلیدی در ارائه سیستم‌های پشتیبانی تصمیم‌گیری قابل تفسیر ایفا کند تا اطمینان حاصل شود که الگوها به درستی شناسایی شده و بیومارکرها به‌طور دقیق اندازه‌گیری می‌شود، که به‌طور مستقیم بر نتایج تأثیر می‌گذارد.

ارزیابی اثربخشی بالینی، مانند بازخورد کاربران، اتکای بالینی و قابلیت تفسیر، باید بهبود یابد و به‌طور دقیق‌تری توصیف شود، به‌ویژه در راهنماهای آینده برای توسعه و گزارش‌دهی مدل (TRIPOD-ML) با وجود اینکه این پروتکل هنوز در حال توسعه است، استاندارد گزارش‌دهی شفاف مدل پیش‌بینی چندمتغیره برای پیش‌آگهی یا تشخیص فردی (TRIPOD:2015) باید مورد توجه قرار گیرد، زیرا راهنمایی‌ها و توصیه‌هایی برای گزارش مدل پیش‌بینی چندمتغیره برای تشخیص یا پیش‌آگهی ارائه می‌دهد. به‌ویژه، نمایندگی باید همیشه مورد توجه قرار گیرد، زیرا یک مفهوم اساسی در کیفیت داده‌ها است که تنوع لازم جمعیت مورد مطالعه را در نسبت متعادل پوشش می‌دهد.

این موضوع به‌ویژه زمانی اهمیت دارد که مدل‌ها قصد دارند رویدادهای دسته‌ای یا دوتایی را در زمینه مشکلات پزشکی پیش‌بینی کنند.

به‌طور خلاصه، برای اینکه مدل‌های پیش‌بینی به درستی عمل کند و نتایج قابل اعتمادی ارائه دهد، باید اطمینان



بیماری‌ها براساس داده‌های به‌دست‌آمده از آزمایش خون استفاده کرد، بدون نیاز به اطلاعات اضافی یا پیچیده.

گزارش‌های مورد بررسی در اینجا از نظر نوع، کیفیت و مقدار منبع داده‌ها متفاوت است و الگوریتم‌های متنوعی از یادگیری ماشین (ML) برای پیش‌بینی نتایج توصیف می‌شود. یافته‌های اصلی نشان می‌دهد که در مطالعات اعتبارسنجی، معیارهای عملکرد مدل‌ها بسیار خوب بوده و همچنین یک تفاوت واضح بین متابولیت‌های استاندارد مرتبط با بیماری و مدل‌های یادگیری ماشین انتخاب شده وجود دارد. این موضوع باعث شده که مدل‌های یادگیری ماشین عملکرد بهتری نسبت به نمرات روش‌های بالینی سنتی داشته باشد.

هرچند هنوز فاصله زیادی بین تحقیقات در زمینه هوش مصنوعی و پیاده‌سازی واقعی آن‌ها در محیط‌های بالینی وجود دارد، این فناوری در حال تحول شیوه‌های معمول پزشکی است.

در این میان، ابزارهای دیجیتال نقش کلیدی در کمک به پزشکان برای ارزیابی شخصی سازی شده بیماران، دارند. این ابزارها به پزشکان امکان می‌دهد تا با استفاده از داده‌های بیمار و فناوری‌های هوشمند، مراقبت‌های بهداشتی را متناسب با نیازهای فردی هر بیمار ارائه دهد.

اگر فقط از داده‌های آزمایش خون معمولی برای ساخت مدل‌های تشخیصی استفاده شود، می‌توان تشخیص‌های پیچیده‌تری را به سرعت و به طور مستقیم در مکان‌های درمانی ارائه داد. این کار نه تنها به تسریع در فرآیند تشخیص کمک می‌کند، بلکه می‌تواند هزینه‌های مربوط به مراقبت‌های بهداشتی را نیز کاهش دهد.

پایان

حاصل شود که داده‌های مورد استفاده نماینده‌ای از جمعیت واقعی و متنوع بیمارانی باشد که در آن‌ها مطالعه می‌شود.

در آینده، استقرار مدل‌های یادگیری ماشین (ML) همچنان با تغییرات داده‌ها در طول زمان مواجه خواهد بود که این موضوع می‌تواند نمایندگی داده‌ها را تحت تأثیر قرار دهد و عملکرد مدل را مختل کند. به عبارت دیگر، تغییرات در داده‌ها به مرور زمان می‌تواند باعث شود که مدل‌ها نتوانند به درستی نماینده جمعیت یا شرایط فعلی باشند و در نتیجه، دقت و کارایی آن‌ها کاهش یابد.

یک مثال بارز به عملکرد مدل‌های COVID-19 اشاره دارد که در طول شیوع اولیه سویه آلفا آموزش دیده‌اند و اکنون با شرایط بیماری جاری، که تحت تأثیر چندین سویه دیگر ویروس قرار دارد، مواجه هستند. همچنین، تأثیر اضافه شدن واکسن‌ها که نتایج بیماری را برای اکثریت افراد مبتلا تغییر داده است، نیز باید در نظر گرفته شود. بنابراین، استقرار باید با مطالعات پایلوت (آزمایشی) جداگانه آغاز شود تا بازخوردی از کارشناسان بهداشت و درمان در مورد تجربه کاربری، رابط کاربری، کارایی و عملکرد ارزیابی در زمان واقعی دریافت شود یعنی قبل از اینکه یک مدل یا ابزار به طور گسترده مورد استفاده قرار گیرد، باید در محیط‌های کوچک و کنترل شده آزمایش شود تا نقاط قوت و ضعف آن شناسایی شود و به بهبود کیفیت آن کمک کند.

نقطه نظر

این مرور به بررسی کاربرد الگوریتم‌های هوش مصنوعی در تشخیص و پیش‌بینی ICD10 با استفاده تنها از آزمایش‌های خون روتین می‌پردازد.

هدف این مطالعه این است که نشان دهد چگونه می‌توان از تکنیک‌های هوش مصنوعی برای تشخیص و پیش‌بینی