

مسئولیت کیفری الگوریتم‌ها و هوش مصنوعی

در تصمیم‌گیری‌های خودکار

(تاریخ دریافت ۱۴۰۵/۰۱/۲۵ تاریخ تصویب ۱۴۰۵/۰۳/۰۴)

مرتضی افشاری

چکیده

گسترش فناوری‌های مبتنی بر هوش مصنوعی و الگوریتم‌های تصمیم‌گیرنده، ساختار بسیاری از فرایندهای اجتماعی، اقتصادی، اداری و قضایی را دستخوش تحول کرده است. امروزه سامانه‌های هوشمند در حوزه‌هایی مانند تشخیص تقلب مالی، ارزیابی اعتباری، تشخیص پزشکی، مدیریت حمل و نقل، نظارت امنیتی و حتی برخی فرآیندهای مرتبط با اجرای قانون، نقش تصمیم‌ساز یا تصمیم‌گیر پیدا کرده‌اند. این تحول اگرچه موجب افزایش سرعت، دقت و کارآمدی شده است، اما پرسش‌های بنیادینی را در حوزه حقوق کیفری مطرح می‌کند؛ از جمله اینکه در صورت وقوع رفتار زیان‌بار یا مجرمانه ناشی از عملکرد یک الگوریتم، چه شخص یا نهادی باید مسئولیت کیفری داشته باشد؟ آیا می‌توان برای خود الگوریتم و هوش مصنوعی شخصیت کیفری قائل شد؟ یا مسئولیت باید متوجه طراحان، توسعه‌دهندگان، مالکان، بهره‌برداران یا نهادهای نظارتی باشد؟ مسئولیت کیفری در نظام‌های حقوقی سنتی بر پایه مفاهیمی همچون اراده، آگاهی، قصد مجرمانه و قابلیت انتساب استوار شده است؛ مفاهیمی که تطبیق آن‌ها با ماهیت غیرانسانی و گاه خودمختار سامانه‌های هوش مصنوعی با دشواری‌هایی همراه است. الگوریتم‌ها فاقد ادراک انسانی، قصد و انگیزه هستند، اما توانایی یادگیری، تغییر رفتار و اتخاذ تصمیم‌هایی را دارند که حتی برای سازندگان آن‌ها قابل پیش‌بینی نیست. از این رو، الگوی سنتی انتساب جرم با چالش جدی مواجه شده است. این مقاله با روش توصیفی

- تحلیلی و با بهره‌گیری از مطالعات حقوقی، به بررسی مبانی مسئولیت کیفری در ارتباط با الگوریتم‌ها و هوش مصنوعی در تصمیم‌گیری‌های خودکار می‌پردازد. یافته‌های پژوهش نشان می‌دهد که در وضعیت کنونی، ایجاد مسئولیت کیفری مستقل برای هوش مصنوعی با اصول بنیادین حقوق کیفری سازگار نیست و راهکار مناسب‌تر، طراحی نظام مسئولیت چندلایه میان اشخاص انسانی و حقوقی مرتبط با چرخه تولید، توسعه، بهره‌برداری و نظارت بر سامانه‌های هوشمند است. همچنین، ضرورت دارد قانون‌گذاران با تدوین مقررات ویژه، معیارهای شفافیت، قابلیت توضیح‌پذیری، کنترل انسانی و نظارت مستمر را به‌عنوان الزامات بهره‌گیری مسئولانه از هوش مصنوعی مورد توجه قرار دهند.

واژگان کلیدی: الگوریتم، هوش مصنوعی، مسئولیت کیفری، تصمیم‌گیری خودکار، جرایم فناوریانه، قابلیت انتساب، حقوق کیفری نوین

مقدمه

تحولات شتابان فناوری اطلاعات و ارتباطات، یکی از مهم‌ترین دگرگونی‌های عصر حاضر را در حوزه تصمیم‌گیری ایجاد کرده است. در گذشته، تصمیم‌های مهم اجتماعی، اقتصادی و حقوقی عمدتاً توسط انسان‌ها اتخاذ می‌شد و مسئولیت ناشی از این تصمیم‌ها نیز به‌طور مستقیم قابل انتساب به اشخاص حقیقی یا حقوقی بود. اما ظهور الگوریتم‌های پیچیده و سامانه‌های هوش مصنوعی، وضعیت جدیدی را پدید آورده است که در آن بخشی از فرایند تصمیم‌گیری از کنترل مستقیم انسان خارج شده و به سیستم‌های محاسباتی واگذار شده است. هوش مصنوعی امروزی صرفاً مجموعه‌ای از دستورهای ثابت و از پیش تعیین‌شده نیست، بلکه بسیاری از سامانه‌های جدید بر پایه یادگیری ماشین و شبکه‌های عصبی مصنوعی فعالیت می‌کنند. این سامانه‌ها قادرند حجم گسترده‌ای از داده‌ها را تحلیل کنند، الگوهای جدید را

شناسایی نمایند و بر اساس تجربه های قبلی عملکرد خود را اصلاح کنند. ویژگی یادگیری و سازگاری این فناوری موجب شده است که رفتار آن ها در برخی موارد غیرقابل پیش بینی باشد. ورود هوش مصنوعی به حوزه هایی که مستقیماً با حقوق و آزادی های افراد ارتباط دارد، اهمیت مسئله مسئولیت را دوچندان کرده است. برای نمونه، اگر سامانه ای هوشمند در فرآیند ارزیابی قضایی یا انتظامی، فردی را به اشتباه خطرناک تشخیص دهد، یا الگوریتمی در حوزه مالی موجب ورود خسارت گسترده به اشخاص شود، تعیین مسئولیت کیفری ناشی از این نتیجه با چالش مواجه خواهد بود. در چنین شرایطی، پرسش اصلی این است که آیا می توان رفتار زیان بار را به یک ابزار فنی نسبت داد یا باید مسئولیت را به انسان های دخیل در طراحی و استفاده از آن بازگرداند. حقوق کیفری کلاسیک بر پایه اصل شخصی بودن مسئولیت کیفری شکل گرفته است. بر اساس این اصل، تنها کسی می تواند مجازات شود که رفتار مجرمانه به او قابل انتساب باشد و میان اراده او و نتیجه زیان بار رابطه ای وجود داشته باشد. در حالی که الگوریتم ها فاقد اراده انسانی و شعور اخلاقی هستند، اما ممکن است تصمیم هایی اتخاذ کنند که آثار مشابه رفتار انسانی داشته باشد. همین مسئله سبب شده است که برخی پژوهشگران از ضرورت بازنگری در مفاهیم سنتی حقوق کیفری سخن بگویند. مسئله مسئولیت کیفری هوش مصنوعی تنها یک موضوع نظری نیست، بلکه با گسترش استفاده عملی از این فناوری، به یکی از چالش های مهم نظام های حقوقی تبدیل شده است. نظام های حقوقی مختلف در حال بررسی این موضوع هستند که چگونه می توان میان توسعه فناوری و حفظ اصول بنیادین عدالت کیفری تعادل برقرار کرد.

بخش اول: کلیات مفهومی و مبانی نظری مسئولیت کیفری الگوریتم‌ها

بند اول: مفهوم الگوریتم و هوش مصنوعی در نظام تصمیم‌گیری خودکار

الگوریتم مجموعه‌ای از دستورها و قواعد منطقی است که برای حل یک مسئله یا انجام یک فرایند مشخص طراحی می‌شود. در ساده‌ترین شکل، الگوریتم‌ها مجموعه‌ای از فرمان‌های از پیش تعیین‌شده هستند که خروجی مشخصی را بر اساس ورودی‌های معین تولید می‌کنند. با توسعه فناوری، الگوریتم‌های سنتی جای خود را به الگوریتم‌های هوشمند داده‌اند که توانایی تحلیل داده‌ها و اصلاح عملکرد خود را دارند. هوش مصنوعی به مجموعه‌ای از فناوری‌ها اطلاق می‌شود که تلاش می‌کنند قابلیت‌هایی مشابه توانایی‌های شناختی انسان مانند یادگیری، استدلال، پیش‌بینی و تصمیم‌گیری را در ماشین‌ها ایجاد کنند. تفاوت اساسی هوش مصنوعی جدید با نرم‌افزارهای سنتی در این است که بسیاری از سامانه‌های هوشمند، به جای پیروی صرف از دستورات ثابت، از داده‌ها می‌آموزند و الگوهای جدید ایجاد می‌کنند. تصمیم‌گیری خودکار زمانی رخ می‌دهد که یک سامانه بدون دخالت مستقیم انسان، بر اساس داده‌ها و مدل‌های محاسباتی، تصمیمی اتخاذ کند. این تصمیم ممکن است صرفاً جنبه پیشنهادی داشته باشد یا به‌طور کامل مبنای اقدام قرار گیرد. هرچه میزان استقلال این سامانه‌ها افزایش یابد، مسئله مسئولیت حقوقی ناشی از عملکرد آن‌ها نیز پیچیده‌تر خواهد شد.

بند دوم: مفهوم مسئولیت کیفری و عناصر تشکیل‌دهنده آن

مسئولیت کیفری در معنای حقوقی، الزام شخص به تحمل آثار و پیامدهای کیفری رفتار مجرمانه‌ای است که به او منتسب می‌شود. تحقق مسئولیت کیفری معمولاً نیازمند وجود سه عنصر اساسی است: رفتار مجرمانه، قابلیت انتساب رفتار به مرتکب و وجود شرایط ذهنی لازم مانند قصد یا تقصیر. در حقوق کیفری، صرف وقوع یک نتیجه زیان‌بار برای ایجاد مسئولیت کافی نیست، بلکه باید مشخص شود چه شخصی با چه میزان از آگاهی و اراده در تحقق آن نقش داشته است. این ویژگی سبب شده است که مسئولیت کیفری عمدتاً متوجه انسان‌ها باشد،

زیرا انسان دارای اراده، قدرت انتخاب و درک ارزش های اخلاقی و اجتماعی است. الگوریتم ها اگرچه می توانند رفتارهایی پیچیده و حتی پیش بینی ناپذیر از خود نشان دهند، اما فاقد عناصر روانی لازم برای تحقق مسئولیت کیفری هستند. آن ها نمی توانند قصد مجرمانه داشته باشند یا مفهوم ممنوعیت قانونی یک رفتار را درک کنند. بنابراین، پذیرش مسئولیت کیفری مستقل برای الگوریتم ها نیازمند بازتعریف اساسی مفاهیم بنیادین حقوق کیفری خواهد بود.

بند سوم: چالش قابلیت انتساب در جرایم ناشی از عملکرد هوش مصنوعی

یکی از مهم ترین چالش های حقوق کیفری در مواجهه با هوش مصنوعی، مسئله انتساب است. در یک سیستم سنتی، رابطه میان عمل انسان و نتیجه مجرمانه معمولاً قابل شناسایی است، اما در سامانه های هوشمند این رابطه ممکن است میان عوامل متعدد توزیع شود. برای مثال، یک تصمیم نادرست توسط یک سامانه هوشمند ممکن است نتیجه اشتباه در طراحی الگوریتم، نقص داده های آموزشی، استفاده نادرست کاربر، فقدان نظارت کافی یا حتی رفتار پیش بینی نشده خود سامانه باشد. در چنین وضعیتی، تعیین اینکه چه کسی باید مسئول شناخته شود، نیازمند تحلیل دقیق نقش هر یک از عوامل دخیل است. این مسئله در الگوریتم های یادگیرنده پیچیده تر می شود؛ زیرا حتی توسعه دهندگان نیز ممکن است نتوانند دقیقاً توضیح دهند که سامانه چگونه به یک نتیجه خاص رسیده است. این ویژگی که به «جعبه سیاه الگوریتمی» مشهور شده، یکی از موانع اصلی در مسیر تعیین مسئولیت کیفری است.

بخش دوم: بیان مسئله و چالش های حقوق کیفری ناشی از تصمیم گیری های خودکار

بند اول: تحول مفهوم فاعل جرم در عصر هوش مصنوعی

حقوق کیفری سنتی بر این پیش فرض اساسی استوار است که جرم، نتیجه رفتار انسانی است و فاعل جرم موجودی دارای اراده، شعور و قدرت انتخاب است. در این چارچوب، مجازات

زمانی مشروعیت پیدا می‌کند که شخص مرتکب، توانایی درک ماهیت رفتار خود و تشخیص ممنوعیت یا مجاز بودن آن را داشته باشد. با ورود هوش مصنوعی به عرصه تصمیم‌گیری، این پیش‌فرض بنیادین با پرسش‌های جدیدی مواجه شده است. سامانه‌های هوشمند امروزی در برخی حوزه‌ها چنان استقلال عملکردی یافته‌اند که خروجی آن‌ها صرفاً انعکاس تصمیم مستقیم یک انسان نیست. یک الگوریتم پیشرفته ممکن است با تحلیل میلیون‌ها داده، تصمیمی اتخاذ کند که حتی طراح آن نیز از نتیجه نهایی آن آگاهی نداشته باشد. این وضعیت، مفهوم سنتی فاعل جرم را با بحران مواجه می‌کند؛ زیرا میان رفتار فنی سامانه و اراده انسانی، فاصله‌ای ایجاد شده است که در بسیاری از موارد تعیین رابطه علی میان آن‌ها دشوار است. با وجود این، پذیرش الگوریتم به عنوان فاعل مستقل جرم با موانع نظری و عملی جدی روبه‌رو است. حقوق کیفری بر مبنای مسئولیت اخلاقی شکل گرفته و مجازات، علاوه بر کارکرد بازدارندگی، دارای جنبه سرزنش اجتماعی است. سرزنش موجودی که فاقد آگاهی، احساس و اراده اخلاقی است، با مبانی فلسفی مجازات سازگار نیست. از این رو، مسئله اصلی نه ایجاد شخصیت کیفری برای الگوریتم، بلکه یافتن سازوکار مناسب برای انتساب مسئولیت به اشخاص انسانی مرتبط با فناوری است.

بند دوم: مسئله تقصیر و پیش‌بینی پذیری رفتار الگوریتم‌ها

یکی از مفاهیم کلیدی در حقوق کیفری، عنصر تقصیر است. تقصیر زمانی تحقق پیدا می‌کند که شخص برخلاف الزامات قانونی و معیارهای متعارف احتیاط رفتار کند و در نتیجه، زمینه وقوع جرم یا ورود خسارت را فراهم آورد. اما در حوزه هوش مصنوعی، تشخیص تقصیر انسانی با پیچیدگی‌های فراوان همراه است. در بسیاری از سامانه‌های هوشمند، رفتار نهایی سیستم حاصل تعامل میان داده‌های آموزشی، معماری الگوریتم، فرآیند یادگیری و شرایط محیطی است. ممکن است توسعه‌دهنده تمام استانداردهای فنی را رعایت کرده باشد، اما سامانه در شرایط خاص تصمیمی غیرمنتظره اتخاذ کند. در چنین شرایطی، انتساب تقصیر به طراح یا بهره‌بردار نیازمند بررسی دقیق میزان کنترل و امکان پیش‌بینی رفتار سیستم است. از سوی دیگر،

نمی توان صرف پیچیدگی فناوری را عاملی برای حذف کامل مسئولیت انسانی دانست. افرادی که سامانه های هوشمند را طراحی، عرضه یا مورد استفاده قرار می دهند، باید نسبت به خطرات احتمالی آن آگاهی داشته باشند و اقدامات لازم برای کاهش آسیب ها را پیش بینی کنند. بنابراین، مفهوم تقصیر در عصر هوش مصنوعی نیازمند بازتعریف بر اساس معیارهایی مانند میزان کنترل، سطح آگاهی فنی، قابلیت پیش بینی و وظیفه نظارت است.

بند سوم: نقش داده ها در شکل گیری مسئولیت کیفری هوش مصنوعی

داده ها مهم ترین عنصر عملکرد بسیاری از سامانه های هوش مصنوعی محسوب می شوند. الگوریتم های یادگیری ماشین برای تصمیم گیری به حجم گسترده ای از اطلاعات نیاز دارند و کیفیت این داده ها تأثیر مستقیمی بر خروجی سیستم دارد. داده های ناقص، نادرست یا تبعیض آمیز می توانند موجب ایجاد تصمیم های ناعادلانه یا حتی نتایج زیان بار شوند. در صورتی که یک سامانه هوشمند بر اساس داده های نادرست آموزش دیده باشد و تصمیمی مجرمانه یا زیان آور اتخاذ کند، این پرسش مطرح می شود که مسئولیت متوجه چه کسی است. آیا تولید کننده داده، توسعه دهنده الگوریتم، مالک سامانه یا استفاده کننده نهایی باید پاسخگو باشد؟ پاسخ به این پرسش نیازمند شناسایی زنجیره مسئولیت در چرخه حیات هوش مصنوعی است. همچنین، مسئله سوگیری الگوریتمی یکی از چالش های مهم حقوقی است. الگوریتم ها ممکن است الگوهای نابرابر موجود در داده های تاریخی را بازتولید کنند و در تصمیم گیری های خود کار موجب تبعیض شوند. اگر چنین تصمیم هایی منجر به نقض حقوق افراد شود، نظام کیفری باید بتواند مرز میان خطای فنی، بی احتیاطی انسانی و رفتار مجرمانه را مشخص کند.

بخش سوم: مرور ادبیات و مطالعات پیشین

بند اول: دیدگاه‌های حقوقی درباره شخصیت و مسئولیت هوش مصنوعی

در ادبیات حقوقی معاصر، دیدگاه‌های متفاوتی درباره جایگاه حقوقی هوش مصنوعی مطرح شده است. گروهی از پژوهشگران معتقدند که با افزایش استقلال سامانه‌های هوشمند، ممکن است در آینده نیاز به ایجاد نوعی شخصیت حقوقی ویژه برای هوش مصنوعی مطرح شود. این دیدگاه با هدف حل مسئله مسئولیت در مواردی ارائه شده است که هیچ‌یک از اشخاص انسانی دخیل، به‌طور مستقیم مقصر شناخته نمی‌شوند. طرفداران این نظریه معتقدند که همان‌گونه که اشخاص حقوقی مانند شرکت‌ها در نظام‌های حقوقی دارای مسئولیت هستند، می‌توان برای سامانه‌های هوشمند نیز نوعی مسئولیت مستقل تعریف کرد. آنان استدلال می‌کنند که پیچیدگی و استقلال عملکردی برخی فناوری‌ها ممکن است ضرورت عبور از مفهوم سنتی مسئولیت انسانی را ایجاد کند. در مقابل، بسیاری از حقوقدانان با این رویکرد مخالف هستند. آنان معتقدند که شخصیت حقوقی شرکت‌ها نیز در نهایت ابزاری برای انتساب رفتار انسان‌هاست و هوش مصنوعی فاقد ویژگی‌هایی مانند آگاهی، اراده و هدف‌مندی است. بنابراین، اعطای شخصیت کیفری به ماشین‌ها می‌تواند موجب تضعیف اصول بنیادین حقوق کیفری شود.

بند دوم: رویکرد مسئولیت انسانی و مدل زنجیره‌ای پاسخگویی

رویکرد غالب در مطالعات حقوقی جدید، بر حفظ محوریت مسئولیت انسانی تأکید دارد. بر اساس این دیدگاه، هوش مصنوعی باید به‌عنوان یک ابزار پیشرفته در نظر گرفته شود و مسئولیت ناشی از عملکرد آن بر اساس میزان نقش و کنترل اشخاص انسانی تعیین شود. در این مدل، تمامی افراد دخیل در چرخه تولید و استفاده از هوش مصنوعی ممکن است در شرایط خاص مسئول شناخته شوند. طراح الگوریتم، مسئول در صورت نقص طراحی یا عدم رعایت استانداردهای لازم، شرکت تولیدکننده در صورت عرضه محصول خطرناک، بهره‌بردار در

صورت استفاده نادرست و نهاد نظارتی در صورت کوتاهی در کنترل سیستم، می توانند موضوع بررسی مسئولیت قرار گیرند. این رویکرد با اصول سنتی حقوق کیفری سازگارتر است، زیرا همچنان عنصر انسانی را به عنوان محور مسئولیت حفظ می کند، اما در عین حال تلاش دارد پاسخ مناسبی برای پیچیدگی های فناوری جدید ارائه دهد.

بند سوم: مطالعات تطبیقی در نظام های حقوقی مختلف

نظام های حقوقی مختلف در سال های اخیر تلاش کرده اند چارچوب هایی برای تنظیم مسئولیت ناشی از هوش مصنوعی ایجاد کنند. برخی نظام ها بر اصل احتیاط و نظارت انسانی تأکید دارند و برخی دیگر به دنبال ایجاد مقررات ویژه برای سامانه های پرخطر هستند. در نظام های حقوقی مبتنی بر حقوق مدنی، توجه بیشتری به تدوین قواعد قانونی مشخص درباره توسعه و استفاده از هوش مصنوعی وجود دارد. در مقابل، نظام های کامن لا بیشتر بر توسعه تدریجی قواعد مسئولیت از طریق رویه قضایی و تفسیر اصول موجود تمرکز دارند. با وجود تفاوت های موجود، نقطه مشترک بسیاری از مطالعات این است که هوش مصنوعی نباید در فضای خلأ حقوقی فعالیت کند. ضرورت ایجاد شفافیت، امکان بررسی تصمیم ها، ثبت فرآیندهای تصمیم گیری و تعیین مسئولیت اشخاص مرتبط، از مهم ترین محورهای مشترک در پژوهش های تطبیقی است.

بخش چهارم: روش شناسی پژوهش

بند اول: روش تحقیق و شیوه گردآوری اطلاعات

این پژوهش با روش توصیفی - تحلیلی انجام شده است. در این روش، ابتدا مفاهیم بنیادی مرتبط با هوش مصنوعی، الگوریتم، تصمیم گیری خودکار و مسئولیت کیفری بررسی شده و سپس با تحلیل مبانی حقوق کیفری، امکان انتساب رفتارهای ناشی از عملکرد سامانه های هوشمند ارزیابی شده است. شیوه گردآوری اطلاعات در این پژوهش کتابخانه ای بوده و

داده‌های مورد نیاز از طریق مطالعه منابع حقوقی، کتاب‌ها، مقالات علمی و مطالعات تطبیقی مرتبط با حقوق فناوری و حقوق کیفری نوین جمع‌آوری شده است.

بند دوم: چارچوب تحلیلی پژوهش

چارچوب تحلیلی این پژوهش بر بررسی رابطه میان سه عنصر اصلی استوار است: نخست، ماهیت فنی و استقلال عملکردی الگوریتم‌ها؛ دوم، اصول بنیادین مسئولیت کیفری؛ و سوم، ضرورت ایجاد تعادل میان توسعه فناوری و حمایت از حقوق افراد. بر اساس این چارچوب، تلاش شده است مشخص شود که آیا مفاهیم سنتی حقوق کیفری توانایی پاسخگویی به چالش‌های ناشی از هوش مصنوعی را دارند یا نیازمند اصلاح و تکمیل هستند.

بخش پنجم: تحلیل مبانی مسئولیت کیفری اشخاص مرتبط با هوش مصنوعی

بند اول: مسئولیت کیفری طراحان و توسعه‌دهندگان الگوریتم‌ها

طراحان و توسعه‌دهندگان سامانه‌های هوش مصنوعی نخستین حلقه در زنجیره ایجاد و عملکرد این فناوری محسوب می‌شوند. آن‌ها با انتخاب معماری الگوریتم، تعیین روش‌های پردازش داده، طراحی مدل‌های یادگیری و تنظیم معیارهای عملکرد، نقش مهمی در شکل‌گیری رفتار نهایی سامانه دارند. از این رو، در مواردی که یک سامانه هوشمند موجب وقوع رفتار مجرمانه یا ورود آسیب جدی شود، بررسی نقش توسعه‌دهندگان از اهمیت ویژه‌ای برخوردار است. مسئولیت کیفری توسعه‌دهندگان را نمی‌توان صرفاً بر مبنای نتیجه زیان‌بار عملکرد الگوریتم تعیین کرد، زیرا فناوری‌های پیچیده همواره با درجه‌ای از عدم قطعیت همراه هستند. برای تحقق مسئولیت کیفری، باید ثابت شود که توسعه‌دهنده از استانداردهای فنی و حرفه‌ای لازم عدول کرده یا با وجود آگاهی از خطرات احتمالی، اقدام مؤثری برای جلوگیری از پیامدهای زیان‌بار انجام نداده است. برای مثال، اگر یک برنامه‌نویس یا تیم توسعه‌دهنده، الگوریتمی را بدون آزمایش‌های لازم در حوزه‌ای حساس مانند تشخیص هویت، امنیت عمومی یا تصمیم‌های مالی عرضه کند و این امر موجب ورود خسارت گسترده شود، ممکن

است رفتار آنان از منظر تقصیر حرفه ای قابل بررسی باشد. در چنین مواردی، مسئولیت کیفری نه به دلیل خطای خود الگوریتم، بلکه به دلیل کوتاهی انسان های مسئول در فرآیند طراحی و ارزیابی فناوری مطرح می شود. با این حال، تعیین مرز میان خطای فنی و تقصیر کیفری یکی از دشوارترین مسائل حقوقی است. توسعه دهندگان نباید صرفاً به دلیل وقوع نتیجه نامطلوب مسئول شناخته شوند، بلکه باید معیارهایی مانند سطح دانش فنی، امکانات موجود، میزان پیش بینی پذیری خطر و استانداردهای زمان طراحی مورد توجه قرار گیرد.

بند دوم: مسئولیت کیفری شرکت ها و اشخاص حقوقی بهره بردار از هوش مصنوعی

بخش قابل توجهی از سامانه های هوش مصنوعی توسط شرکت ها، سازمان ها و مؤسسات بزرگ مورد استفاده قرار می گیرند. این اشخاص حقوقی ممکن است از الگوریتم ها برای تصمیم گیری در حوزه هایی مانند استخدام، اعتبارسنجی مالی، تبلیغات، امنیت یا مدیریت منابع انسانی استفاده کنند. در چنین شرایطی، اگر استفاده نادرست یا غیرمسئولانه از فناوری موجب وقوع جرم شود، مسئولیت شخص حقوقی اهمیت پیدا می کند. در نظام های حقوق کیفری جدید، مسئولیت کیفری اشخاص حقوقی به عنوان یک استثنا بر اصل سنتی مسئولیت فردی پذیرفته شده است. دلیل پذیرش این نوع مسئولیت آن است که بسیاری از تصمیم های مهم در سازمان های بزرگ حاصل عملکرد یک فرد خاص نیست، بلکه نتیجه سیاست ها، ساختار مدیریتی و فرآیندهای سازمانی است. در حوزه هوش مصنوعی، شرکت ها نمی توانند صرفاً با استناد به پیچیدگی فناوری، خود را از مسئولیت مبرا کنند. استفاده کنندگان حرفه ای از سامانه های هوشمند باید پیش از بهره برداری، میزان خطر، قابلیت اعتماد، نحوه پردازش داده ها و امکان کنترل نتایج را ارزیابی کنند. بنابراین، چنانچه یک سازمان بدون رعایت الزامات ایمنی و نظارتی، تصمیم گیری های حساس را به الگوریتم واگذار کند و این امر موجب نقض حقوق افراد شود، امکان بررسی مسئولیت کیفری مدیران یا شخص حقوقی وجود خواهد داشت.

بند سوم: مسئولیت کاربران و بهره‌برداران نهایی سامانه‌های هوشمند

اگرچه طراحی الگوریتم نقش اساسی در عملکرد هوش مصنوعی دارد، اما کاربران نهایی نیز در بسیاری از موارد دارای نقش تعیین‌کننده هستند. یک سامانه هوشمند ممکن است به‌درستی طراحی شده باشد، اما نحوه استفاده نادرست از آن موجب ایجاد نتیجه مجرمانه شود. برای نمونه، کاربری که آگاهانه داده‌های نادرست وارد یک سامانه تصمیم‌گیری می‌کند یا از خروجی الگوریتم برای تبعیض، فریب یا آسیب به دیگران استفاده می‌کند، نمی‌تواند مسئولیت خود را به فناوری منتقل کند. در این وضعیت، هوش مصنوعی صرفاً ابزار ارتکاب جرم محسوب می‌شود و مسئولیت متوجه شخص استفاده‌کننده خواهد بود. از منظر حقوق کیفری، استفاده از فناوری‌های جدید نباید موجب ایجاد مصونیت برای مرتکبان شود. همان‌گونه که استفاده از رایانه یا ابزارهای ارتباطی، مسئولیت کیفری انسان را از بین نمی‌برد، بهره‌گیری از هوش مصنوعی نیز نمی‌تواند به‌عنوان عاملی برای حذف مسئولیت تلقی شود.

بخش ششم: راهکارهای حقوقی و تقنینی برای تنظیم مسئولیت کیفری هوش مصنوعی

بند اول: ضرورت تدوین قوانین اختصاصی در حوزه هوش مصنوعی

سرعت توسعه فناوری هوش مصنوعی موجب شده است بسیاری از نظام‌های حقوقی با خلأهای قانونی مواجه شوند. قوانین کیفری موجود در بسیاری از کشورها برای شرایطی طراحی شده‌اند که در آن رفتار مجرمانه مستقیماً توسط انسان انجام می‌شود، در حالی که تصمیم‌گیری‌های خودکار وضعیت متفاوتی ایجاد کرده‌اند. ضرورت تدوین قوانین اختصاصی در این حوزه به معنای کنار گذاشتن اصول سنتی حقوق کیفری نیست، بلکه هدف آن تکمیل این اصول متناسب با شرایط جدید است. قانون‌گذار باید معیارهای مشخصی برای تعیین مسئولیت اشخاص دخیل در چرخه تولید و استفاده از هوش مصنوعی ایجاد کند. این قوانین باید

موضوعاتی مانند سطح استقلال سامانه، میزان خطر، وظایف نظارتی، الزامات ثبت داده ها، امکان بررسی تصمیم ها و مسئولیت ناشی از نقص های فنی را مورد توجه قرار دهند.

بند دوم: ایجاد اصل نظارت و کنترل انسانی بر تصمیم های مهم

یکی از مهم ترین راهکارهای جلوگیری از خلأ مسئولیت، الزام به حفظ نقش انسان در تصمیم های حساس است. هرچند هوش مصنوعی می تواند سرعت و دقت تصمیم گیری را افزایش دهد، اما در حوزه هایی که با حقوق اساسی افراد ارتباط دارد، نباید جایگزین کامل انسان شود. اصل کنترل انسانی اقتضا می کند که تصمیم های مهم مانند تصمیم های قضایی، کیفری، استخدامی یا محدود کننده حقوق افراد، امکان بررسی و مداخله انسانی داشته باشند. وجود ناظر انسانی می تواند علاوه بر کاهش خطا، مسیر انتساب مسئولیت را نیز روشن تر کند. این اصل همچنین مانع از آن می شود که سازمان ها مسئولیت تصمیم های خود را به یک سیستم فنی غیر قابل پاسخگو منتقل کنند. مسئولیت پذیری انسانی باید به عنوان یکی از اصول اساسی حکمرانی هوش مصنوعی پذیرفته شود.

بند سوم: ضرورت شفافیت و توضیح پذیری الگوریتم ها

یکی از مشکلات مهم سامانه های هوشمند، عدم شفافیت برخی الگوریتم های پیچیده است. در بسیاری از مدل های یادگیری عمیق، حتی متخصصان نیز نمی توانند به سادگی توضیح دهند که چرا سیستم به نتیجه خاصی رسیده است. این ویژگی می تواند با اصول دادرسی عادلانه و حق اعتراض افراد تعارض پیدا کند. فردی که بر اساس تصمیم یک الگوریتم از حقی محروم شده است، باید بتواند دلیل اتخاذ آن تصمیم را بداند و امکان اعتراض داشته باشد. از این رو، توسعه الگوریتم های توضیح پذیر و ایجاد الزام قانونی برای ثبت فرآیند تصمیم گیری، از مهم ترین اقدامات در جهت افزایش پاسخگویی سامانه های هوشمند است.

بند چهارم: ایجاد نظام ارزیابی خطر برای سامانه‌های هوش مصنوعی

همه سامانه‌های هوش مصنوعی دارای سطح یکسانی از خطر نیستند. استفاده از یک الگوریتم برای پیشنهاد فیلم یا موسیقی، تفاوت اساسی با استفاده از آن برای تصمیم‌های قضایی، امنیتی یا پزشکی دارد. بنابراین، نظام حقوقی باید میان کاربردهای مختلف هوش مصنوعی تفکیک قائل شود و برای سامانه‌های پرخطر، الزامات سخت‌گیرانه‌تری تعیین کند. ارزیابی پیشینی خطر، آزمایش‌های دوره‌ای، نظارت مستقل و گزارش‌دهی شفاف می‌تواند از وقوع بسیاری از آسیب‌ها جلوگیری کند.

نتیجه‌گیری

گسترش هوش مصنوعی و الگوریتم‌های تصمیم‌گیرنده، یکی از مهم‌ترین چالش‌های حقوق کیفری معاصر را ایجاد کرده است. این فناوری‌ها اگرچه فرصت‌های فراوانی برای افزایش کارآمدی و کاهش خطا فراهم کرده‌اند، اما در عین حال امکان بروز آسیب‌ها و رفتارهای غیرقابل پیش‌بینی را نیز افزایش داده‌اند. بررسی مبانی حقوق کیفری نشان می‌دهد که پذیرش مسئولیت کیفری مستقل برای الگوریتم‌ها و هوش مصنوعی در شرایط فعلی با اصول بنیادین این حوزه سازگار نیست. الگوریتم فاقد اراده، آگاهی و قابلیت سرزنش اخلاقی است و نمی‌توان همانند انسان آن را موضوع مجازات کیفری قرار داد. با این حال، عدم پذیرش مسئولیت مستقل هوش مصنوعی به معنای ایجاد خلأ مسئولیتی نیست. راهکار مناسب، ایجاد نظام مسئولیت چندلایه است که در آن نقش تمامی اشخاص دخیل در چرخه تولید، توسعه، عرضه، استفاده و نظارت بر سامانه‌های هوشمند مورد ارزیابی قرار گیرد. در این نظام، توسعه‌دهندگان باید نسبت به طراحی ایمن و استاندارد الگوریتم‌ها مسئول باشند، شرکت‌ها باید خطرات استفاده از فناوری را مدیریت کنند و کاربران نیز باید در بهره‌برداری صحیح از سامانه‌های هوشمند دقت لازم را داشته باشند. آینده حقوق کیفری در مواجهه با هوش مصنوعی نیازمند ایجاد تعادل میان دو هدف اساسی است: نخست، حمایت از نوآوری و پیشرفت فناوری

و دوم، تضمین حقوق و آزادی‌های افراد. دستیابی به این هدف مستلزم تدوین قوانین دقیق، افزایش شفافیت الگوریتمی، حفظ نظارت انسانی و بازنگری مستمر در مفاهیم سنتی مسئولیت کیفری است.

منابع و مآخذ

۱. آشوری، محمد. (۱۴۰۰). حقوق جزای عمومی. تهران: انتشارات سمت.
۲. اردبیلی، محمدعلی. (۱۴۰۲). حقوق جزای عمومی، جلد اول. تهران: نشر میزان.
۳. باقری اصل، علیرضا. (۱۴۰۱). حقوق فناوری اطلاعات و ارتباطات. تهران: انتشارات خرسندی.
۴. جعفری لنگرودی، محمدجعفر. (۱۳۹۹). ترمینولوژی حقوق. تهران: گنج دانش.
۵. خزائی، محمدرضا. (۱۴۰۱). حقوق کیفری فناوری‌های نوین. تهران: نشر جنگل.
۶. صادقی، حسین. (۱۴۰۰). مسئولیت کیفری اشخاص حقوقی در حقوق ایران. تهران: انتشارات دانشگاه تهران.
۷. کاتوزیان، ناصر. (۱۳۹۸). فلسفه حقوق و مبانی مسئولیت. تهران: شرکت سهامی انتشار.
۸. Balkin, J. M. (۲۰۱۵). *The Path of Robotics Law*. California Law Review Circuit.
۹. Bostrom, N. (۲۰۱۴). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.
۱۰. Bryson, J. J. (۲۰۱۸). *Patience Is Not Necessary for Responsibility*. AI and Society.
۱۱. Calo, R. (۲۰۱۵). *Robotics and the Lessons of Cyberlaw*. California Law Review.

۱۲. Crawford, K. (۲۰۲۱). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.
۱۳. Floridi, L., & Sanders, J. W. (۲۰۰۴). *On the Morality of Artificial Agents*. *Minds and Machines*.
۱۴. Gless, S., Silverman, E., & Weigend, T. (۲۰۱۶). *If Robots Cause Harm, Who Is to Blame? Self-driving Cars and Criminal Liability*. *New Criminal Law Review*.
۱۵. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (۲۰۱۶). *The Ethics of Algorithms: Mapping the Debate*. *Big Data & Society*.
۱۶. Russell, S., & Norvig, P. (۲۰۲۱). *Artificial Intelligence: A Modern Approach*. Pearson.