

بازتعریف رکن معنوی جرم در بستر کنش‌های الگوریتمی و سامانه‌های خودآموز

هدیه افروزی (نویسنده مسئول)

دانشجوی کارشناسی ارشد، گروه حقوق جزا و جرم‌شناسی، واحد لاهیجان، دانشگاه آزاد اسلامی، لاهیجان، ایران.

hedeyeh.afrooz@iaui.ac.ir

هایده شیرزاد راجعونی

دانشجوی کارشناسی ارشد، گروه حقوق جزا و جرم‌شناسی، واحد لاهیجان، دانشگاه آزاد اسلامی، لاهیجان، ایران.

hayedeh.shirzad@iaui.ac.ir

یاسمن رمضان مقدم واجاری

دانشجوی کارشناسی ارشد، گروه حقوق جزا و جرم‌شناسی، واحد لاهیجان، دانشگاه آزاد اسلامی، لاهیجان، ایران.

y.moghaddam@iaui.ac.ir

امیررضا محمودی

استادیار، گروه حقوق جزا و جرم‌شناسی، واحد لاهیجان، دانشگاه آزاد اسلامی، لاهیجان، ایران.

Dr.mahmoudi@iaui.ac.ir

چکیده

تحول سامانه‌های هوشمند و الگوریتم‌های خودآموز، الگوهای سنتی انتساب کیفری را با چالش‌های بنیادین مواجه ساخته است؛ زیرا در این بستر، رفتار مجرمانه یا زیان‌بار همیشه محصول کنش مستقیم و هم‌زمان یک فاعل انسانی نیست، بلکه می‌تواند از برهم‌کنش میان طراحی، آموزش، استقرار و عملکرد سامانه‌هایی ناشی شود که از استقلال عملیاتی نسبی برخوردارند. مسئله اصلی این پژوهش، بازتعریف رکن معنوی جرم در مواجهه با کنش‌های الگوریتمی و تبیین معیارهای انتساب قصد، آگاهی و اراده در این بستر است. پژوهش حاضر با روش توصیفی - تحلیلی و با بهره‌گیری از منابع حقوق کیفری، فقهی و میان‌رشته‌ای، ظرفیت مفاهیم سنتی حقوق کیفری را در مواجهه با این تحولات بررسی می‌کند. یافته‌ها نشان می‌دهد که سامانه‌های هوشمند، با وجود عاملیت فنی و نقش مؤثر در تولید نتایج، به دلیل فقدان اراده آزاد، آگاهی هنجاری و قابلیت سرزنش کیفری، واجد فاعلیت و مسئولیت کیفری مستقل نیستند. بنابراین، احراز رکن معنوی باید بر اساس رفتار و نقش کنشگران انسانی و با توجه به معیارهایی همچون کنترل مؤثر بر سامانه، قابلیت پیش‌بینی خطر، امکان مداخله و جلوگیری از نتیجه، رعایت الزامات احتیاطی و میزان شفافیت و توضیح‌پذیری سامانه صورت گیرد. همچنین، مفاهیمی مانند تسبیب الگوریتمی، تقصیر ساختاری، مسئولیت توزیعی، تقصیر مقدم و مسئولیت نیابتی می‌توانند چارچوبی مناسب برای تحلیل انتساب کیفری در جرایم ناشی از عملکرد سامانه‌های خودآموز فراهم آورند. بر این اساس، بازتعریف رکن معنوی مستلزم گذار از الگوی فردمحور به رویکردی چندلایه و ساختارمحور است که با اقتضائات حقوق کیفری ژورنالی و چالش‌های فناوری نوین سازگار باشد.

کلیدواژه‌ها: رکن معنوی جرم؛ کنش‌های الگوریتمی؛ سامانه‌های خودآموز؛ عاملیت الگوریتمی؛ انتساب کیفری.

مقدمه:

گسترش شتابان هوش مصنوعی، یادگیری ماشین و سامانه‌های خودآموز، صرفاً ابزارهای ارتکاب یا تسهیل جرم را دگرگون نکرده، بلکه مبانی تحلیل مسئولیت کیفری را نیز در معرض بازاندیشی قرار داده است. در الگوی کلاسیک حقوق کیفری، جرم به رفتار انسانی آگاهانه و ارادی منتسب می‌شود و رکن معنوی، پیوند ذهنی میان فاعل، رفتار مجرمانه و نتیجه حاصل از آن را آشکار می‌سازد. قصد، علم، پذیرش خطر، بی‌احتیاطی و بی‌مبالاتی، همگی بر این پیش‌فرض استوارند که شخص انسانی از توان ادراک، انتخاب و کنترل رفتار خود برخوردار است. با این حال، ورود سامانه‌هایی که بر مبنای داده‌های گسترده تصمیم‌سازی می‌کنند، از نتایج پیشین می‌آموزند و در برخی موقعیت‌ها بدون مداخله لحظه‌ای انسان عمل می‌کنند، کارآمدی این الگوی سنتی را با تردید مواجه ساخته است.

در بسیاری از رخدادهای مرتبط با سامانه‌های هوشمند، نتیجه زیان‌بار یا مجرمانه، حاصل تصمیم آنی یک شخص واحد نیست؛ بلکه در زنجیره‌ای از طراحی نرم‌افزار، انتخاب و پالایش داده‌ها، تعیین اهداف و محدودیت‌ها، آموزش مدل، استقرار سامانه، نظارت فنی و بهره‌برداری عملیاتی شکل می‌گیرد. در چنین وضعیتی، ممکن است تصمیم نهایی به‌وسیله سامانه اتخاذ یا اجرا شود، اما زمینه‌های خطر، قابلیت‌های انحرافی و امکان وقوع نتیجه، در مراحل پیشینی و به‌واسطه تصمیم‌ها یا ترک‌فعل‌های انسانی فراهم آمده باشد. در این چارچوب، پرسش کانونی این است که چگونه می‌توان در غیاب فاعلیت مستقل سامانه، رکن معنوی را به کنشگران انسانی در لایه‌های پیشینی انتساب داد؟

این وضعیت، پدیده‌ای موسوم به «شکاف مسئولیت» را پدید می‌آورد؛ بدین معنا که میان وقوع یک نتیجه زیان‌بار و امکان انتساب آن به شخصی معین، فاصله‌ای ایجاد می‌شود. پیچیدگی، عدم شفافیت، خودمختاری عملیاتی نسبی و قابلیت تحول سامانه‌های خودآموز، احراز عناصر سنتی سوءنیت و تقصیر را دشوارتر می‌سازد. برای نمونه، در مواردی که خروجی سامانه برخلاف پیش‌بینی اولیه عمل می‌کند، باید روشن شود آیا خطر مزبور برای اشخاص دخیل قابل پیش‌بینی بوده است؛ آنان تا چه اندازه بر عملکرد سامانه کنترل مؤثر داشته‌اند؛ آیا امکان مداخله، توقف یا اصلاح تصمیم وجود داشته است؛ و آیا تکالیف حرفه‌ای و احتیاطی لازم در طراحی، آموزش و نظارت رعایت شده است یا خیر.

بر این مبنا، بازتعریف رکن معنوی در بستر کنش‌های الگوریتمی، به معنای انتساب اراده یا قصد انسانی به ماشین نیست؛ بلکه مستلزم بازسازی معیارهای احراز وضعیت ذهنی کنشگران انسانی در یک ساختار چندلایه است. در این چارچوب، قصد و آگاهی صرفاً به لحظه وقوع نتیجه محدود نمی‌شود، بلکه می‌تواند در علم به خطرهای قابل انتظار، پذیرش آگاهانه ریسک، بی‌اعتنایی به هشدارهای فنی، فقدان نظارت مؤثر، طراحی نایمن یا استمرار بهره‌برداری از سامانه معیوب ظهور یابد. بنابراین، تحلیل مسئولیت کیفری باید از الگوی خطیِ مباشر و نتیجه فاصله گرفته و به بررسی نقش‌های پیشینی، ساختاری و واسطه‌ای در تولید وضعیت خطرناک توجه کند.

پژوهش حاضر با روش توصیفی - تحلیلی، در پی پاسخ به این پرسش است که رکن معنوی جرم در جرایم ناشی از کنش‌های الگوریتمی و عملکرد سامانه‌های خودآموز بر چه مبنایی قابل احراز و انتساب است. فرض اصلی آن است که سامانه هوشمند، هرچند واجد عاملیت فنی است، فاقد فاعلیت کیفری مستقل باقی می‌ماند و مسئولیت کیفری باید با معیارهایی چون کنترل مؤثر، قابلیت پیش‌بینی خطر، امکان مداخله و مهار، شفافیت و توضیح‌پذیری، و رعایت تکالیف احتیاطی به کنشگران انسانی مرتبط شود. در این راستا، ظرفیت مفاهیمی چون تسبیب الگوریتمی، تقصیر ساختاری، تقصیر مقدم و مسئولیت توزیعی برای بازفهم رکن معنوی و پرهیز از خلأ مسئولیت بررسی می‌شود.

۱. تحلیل رکن معنوی در بستر کنش‌های الگوریتمی با تکیه بر معیارهای فاعلیت کیفری

در تحلیل رکن معنوی جرم، نقطه عزیمت همچنان بر رابطه ذهنی میان فاعل و رفتار مجرمانه استوار است؛ با این حال، در محیط‌های الگوریتمی و سامانه‌های خودآموز، این رابطه دیگر به سادگی از مسیر فاعل انسانی مستقیم قابل شناسایی نیست. از این رو، بررسی قصد، آگاهی و اراده باید از سطح رفتار ظاهری فراتر رود و به سطح کنترل، طراحی، آموزش و جهت‌دهی سامانه نیز تسری یابد، زیرا در چنین ساختاری، ممکن است نتیجه زیان‌بار از دل فرایندی برآید که هیچ اراده کیفری مستقل انسانی در لحظه اجرای آن حضور ندارد (قیاسی و محترم‌قلاتی، ۱۳۹۶، ۶۱۱ و ۶۳۱؛ رضاپور، ۱۴۰۴، ۲-۱). این وضعیت، نقطه عزیمت دکترین حقوق کیفری برای عبور از بن‌بست «عاملیت فنی» است. چرا که در ادبیات نوین حقوقی، پذیرش فقدان فاعلیت مستقل کیفری در سامانه‌های هوشمند، به یک پیش‌فرض مقدماتی بدل شده است. از این رو، ارزش افزوده تحلیل حقوقی نه در اثبات مکرر «فقدان اراده آزاد در ماشین»، بلکه در بازتعریف معیارهای انتساب هنجاری رکن معنوی به عامل انسانی است. زمانی که الگوریتم خودآموز به دلیل یادگیری عمیق، رفتاری منتهی به نقض قاعده کیفری بروز می‌دهد، عنصر روانی باید از «روان‌شناختی‌گرایی لحظه‌ای» به «تعهدات هنجاری در چرخه حیات سامانه» ارتقا یابد. در این الگو، انتساب عنصر روانی بر سه شاخص عینی استوار است: (۱) کنترل مؤثر بر قلمرو ریسک؛ (۲) پیش‌بینی‌پذیری هنجاری انحرافات الگوریتمی؛ و (۳) قصور ارادی در طراحی و نظارت تدافعی. بدین سان، عاملیت فنی سامانه، مانع انتساب نیست، بلکه «مجرای انتقال تقصیر انسانی» محسوب می‌شود.

صرف تحقق رفتار مادی یا حتی وقوع نتیجه زیان‌بار برای انتساب مسئولیت کیفری کفایت نمی‌کند، بلکه باید رابطه ذهنی مرتکب با عناصر مادی جرم نیز احراز شود. به بیان دیگر، تحقق مسئولیت کیفری در گرو آن است که قصد و آگاهی نسبت به رفتار مجرمانه و عناصر سازنده آن قابل اثبات باشد. از همین رو، هرگاه قانون‌گذار نوع خاصی از جهت‌گیری ذهنی را به صراحت بیان نکرده باشد، اصل بر آن است که آگاهی یا قصد نسبت به اجزای رفتار مجرمانه باید مورد بررسی قرار گیرد و بدون چنین احرازی، انتساب کامل رکن معنوی با دشواری مواجه خواهد بود. این مسئله در جایی پیچیده‌تر می‌شود که رفتار مجرمانه نه به صورت مستقیم، بلکه از طریق واسطه‌های انسانی، سازمانی یا الگوریتمی تحقق پیدا کند. در چنین وضعیتی، محور اصلی تحلیل دیگر صرفاً انجام‌دهنده ظاهری رفتار نیست، بلکه معیار اصلی «کنترل بر رفتار» است؛ یعنی باید احراز شود که چه کسی با وجود عدم مباشرت فیزیکی، در موقعیتی قرار داشته که بتواند مسیر رفتار را هدایت، تنظیم یا متوقف کند (قربانی و سبحانی، ۱۴۰۳، ۱۲۴ و ۱۲۵).

در بستر سامانه‌های خودآموز، این بحث به مسئله «کنترل بر سازمان» پیوند می‌خورد. در اینجا، مسئولیت کیفری تنها با اتکا به اجرای نهایی رفتار قابل تبیین نیست، بلکه باید بررسی شود چه کسی بر ساختاری که رفتار از دل آن صادر شده تسلط مؤثر داشته است. چنین تسلطی زمانی احراز می‌شود که اجزای فنی ساختار تابع جهت‌دهی مدیریتی باشند و امکان جایگزینی یا اصلاح اجزا برای استمرار عملکرد وجود داشته باشد. این قابلیت نشان می‌دهد که نقش اصلی در کنترل فرایند مجرمانه با کسی است که سامانه را در وضعیتی نگه می‌دارد که نتیجه مجرمانه بتواند از طریق آن تحقق یابد (قربانی و سبحانی، ۱۴۰۳، ۱۲۱). بنابراین، میان «دستوردهی» و «کنترل سازمانی» باید تفاوت گذاشت؛ آنچه در اینجا اهمیت دارد نه فقط صدور یک دستور مشخص، بلکه امکان حفظ و هدایت فرایندی است که در نهایت به وقوع جرم می‌انجامد. هرگاه اراده مباشر یا خروجی سامانه در سطره ساختار عامل هدایت‌کننده قرار گیرد، رکن معنوی را باید در سطحی جست‌وجو کرد که کنترل مؤثر بر مسیر وقوع نتیجه در آن قرار دارد (قربانی و سبحانی، ۱۴۰۳، ۱۲۳-۱۲۰).

با وجود این پیچیدگی‌ها، رفتار ناشی از خطای الگوریتمی را نباید به طور مطلق به خود سامانه نسبت داد؛ زیرا نظام کیفری بر پایه مسئولیت عوامل واجد شعور حقوقی استوار است. از این رو، تحلیل فاعلیت باید به جای تمرکز بر رفتار مادی مباشر، بر کنشگران انسانی تمرکز یابد که در لایه‌های پیشینی نظیر طراحی، تنظیم، آموزش داده‌ها یا بهره‌برداری نقش داشته‌اند. در این بستر، مفهوم سنتی مباشرت دستخوش دگرگونی شده و فاعلیت انسانی در قالب «فاعل بالتبع» یا «فاعل بالقوه» بازتعریف می‌شود؛ بدین معنا که شخص با ایجاد یا هدایت بستر فناورانه، زمینه تولید کنش زیان‌بار بعدی را به طور غیرمستقیم فراهم می‌سازد (ملکی، ۱۴۰۴، ۱۶۶-۱۶۵).

۲. بازتعریف قصد، آگاهی و اراده در زنجیره کنش‌های هوشمند و پیش‌نگر

با ورود فناوری‌های هوشمند به فرایند کیفری، مسئله رکن معنوی ابعاد تازه‌ای پیدا کرده است. عدالت کیفری دیگر تنها در مقام واکنش به جرم پس از وقوع آن عمل نمی‌کند، بلکه به تدریج به سوی الگوهای پیش‌نگر، داده‌محور و مبتنی بر تحلیل الگوهای خطر حرکت کرده است. در این چارچوب، سامانه‌های هوشمند نه فقط ابزار کشف جرم، بلکه بخشی از سازوکار حکمرانی کیفری محسوب می‌شوند و در شناسایی زمینه‌های خطر، اولویت‌بندی مداخله و حتی شکل‌دهی به تصمیم‌های نظارتی نقش دارند. چنین وضعیتی موجب می‌شود که تحلیل مسئولیت کیفری و به‌ویژه رکن معنوی، دیگر منحصر به اراده فردی نباشد و سهم ساختارهای فناورانه در شکل‌دهی، تسهیل یا تقویت رفتارهای زیان‌بار نیز مورد توجه قرار گیرد (احمدی خونسارکی و همکاران، ۱۴۰۵، ۶).

در بستر حکمرانی الگوریتمی، اراده انسانی همواره در خلأ شکل نمی‌گیرد، بلکه ممکن است در معرض سازوکارهایی قرار گیرد که از طریق تحلیل داده‌های رفتاری، شخصی‌سازی محیط تصمیم و جهت‌دهی نامحسوس ترجیحات، دامنه انتخاب و کیفیت داوری فرد را تحت تأثیر قرار می‌دهند. بر این اساس، احراز قصد و آگاهی در کنش‌های مرتبط با سامانه‌های هوشمند نباید صرفاً بر رفتار نهایی یا انتخاب ظاهری شخص مبتنی باشد، بلکه باید فرایند شکل‌گیری تصمیم، میزان استقلال اراده، امکان مقاومت در برابر هدایت فناورانه و سطح اثرگذاری ساختار الگوریتمی بر تصمیم نیز ارزیابی شود (انصاری و مهدانیان، ۱۴۰۴، ۲۳۹-۲۴۰).

۱-۲. حکمرانی پیش‌نگر و گذار به پیشگیری داده‌محور

تحول در روش‌های کشف و ردیابی جرایم، مفهوم سنتی رکن معنوی را با چالش‌های جدیدی مواجه کرده است. با بهره‌گیری از هوش مصنوعی و یادگیری عمیق، امکان رصد داده‌های سنجش از دور، حسگرهای محیطی و کلان‌داده‌ها فراهم شده است که خروجی آن استخراج الگوهای پیچیده و غیرخطی از رفتارهای پرخطر است. این تحول، واکنش کیفری را از حالت پسینی به رویکردی پیش‌بینانه و هشداردهنده سوق می‌دهد؛ به طوری که جرم دیگر یک واقعه تک‌عاملی تلقی نمی‌شود، بلکه فرایندی پویا و رصدپذیر در بستر زمان به شمار می‌رود (احمدی خونسارکی و همکاران، ۱۴۰۵، ۸). با این حال، از منظر حقوق کیفری، پیش‌بینی الگوریتمی خطر وقوع جرم نباید با قصد مجرمانه یا علم پیشینی عامل انسانی یکسان انگاشته شود، چرا که تحلیل خروجی‌های محاسباتی برای انتساب ذهنی جرم، نیازمند اثبات اراده معطوف به نتیجه در لایه‌های تصمیم‌گیری انسانی است (احمدی خونسارکی و همکاران، ۱۴۰۵، ۸؛ رضاپور، ۱۴۰۴، ۲).

۲-۲. شفافیت الگوریتمی و پیش‌شرط‌های ارزیابی آگاهی

گسترش اتکای نظام کیفری به الگوریتم‌ها و داده‌های کلان، بدون شفافیت کافی، خود می‌تواند به دشواری در انتساب مسئولیت بینجامد. هرچه سامانه از حیث شیوه تصمیم‌گیری پیچیده‌تر و غیرشفاف‌تر باشد، تشخیص اینکه نتیجه زیان‌بار به کدام یک از

عوامل انسانی اعم از طراح، بهره‌بردار، ناظر یا تصمیم‌گیرنده قابل انتساب است، دشوارتر خواهد شد. به همین دلیل، شفافیت الگوریتمی و قابلیت توضیح‌پذیری صرفاً یک مطالبه فنی نیست، بلکه پیش‌شرطی اساسی برای ارزیابی حقوقی رکن معنوی به شمار می‌آید؛ زیرا بدون فهم سازوکار تصمیم‌سازی سامانه، احراز آگاهی، پیش‌بینی‌پذیری و میزان پذیرش خطر از سوی عامل انسانی با ابهام جدی روبه‌رو می‌شود (احمدی خونسارکی و همکاران، ۱۴۰۵، ۸).

شفافیت در این حوزه، صرفاً به افشای فنی سازوکار سامانه محدود نمی‌شود، بلکه باید امکان فهم، ارزیابی و بازنگری انسانی تصمیم را نیز فراهم سازد. هرگاه منطق مؤثر در تولید خروجی، هشدارهای مربوط به خطر یا مبنای داده‌ای تصمیم از دسترس اشخاص مسئول خارج باشد، تشخیص میزان علم آنان نسبت به احتمال وقوع نتیجه زیان‌بار با دشواری روبه‌رو خواهد شد. بنابراین، قابلیت توضیح‌پذیری و امکان بازبینی انسانی، افزون بر کارکرد تضمینی خود، معیاری برای سنجش دسترسی شخص به اطلاعات خطر و امکان احراز آگاهی یا بی‌مبالاتی وی در فرایند انتساب کیفری به شمار می‌آید (تنگستانی و مرادی برلیان، ۱۴۰۴، ۲۸ و ۴۹).

اهمیت شفافیت از آن جهت دوچندان می‌شود که در محیط‌های داده‌محور، الگوریتم‌ها صرفاً به پردازش ختای اطلاعات بسنده نمی‌کنند، بلکه می‌توانند در قالب سازوکارهای پیش‌بینی و شخصی‌سازی، بر ادراک، انتخاب و رفتار اشخاص نیز اثر بگذارند. از این‌رو، هرچه سازوکار اثرگذاری سامانه پنهان‌تر و غیرقابل فهم‌تر باشد، ارزیابی میزان آگاهی اشخاص انسانی نسبت به خطرهای ناشی از آن و نیز سنجش استقلال تصمیم‌گیری آنان با دشواری بیشتری روبه‌رو خواهد شد (انصاری و مهدانیان، ۱۴۰۴، ۲۳۷ و ۲۳۹).

در این میان، معیار بی‌پروایی ذهنی در حقوق کیفری تطبیقی بر پیش‌بینی یک خطر ناموجه استوار است و از این جهت، می‌تواند در ارزیابی رفتار طراح، بهره‌بردار یا ناظر سامانه‌های هوشمند مورد استفاده قرار گیرد؛ با این توضیح که صرف پیش‌بینی‌پذیری فنی خروجی سامانه، معادل با قصد مجرمانه یا علم پیشینی عامل انسانی نیست و برای انتساب رکن معنوی، باید اراده معطوف به نتیجه در لایه‌های تصمیم‌گیری انسانی احراز شود. هرچه سامانه از حیث سازوکار تصمیم‌سازی پیچیده‌تر و غیرشفاف‌تر باشد، تشخیص میزان آگاهی و پذیرش خطر از سوی اشخاص دخیل دشوارتر خواهد بود و در نتیجه، شفافیت الگوریتمی به شرطی ضروری برای ارزیابی حقوقی رکن معنوی بدل می‌شود (نحوی و همکاران، ۱۴۰۱، ۸ و ۱۱).

همین ابهام در موارد مربوط به نقص فنی، اختلال حسگرها، خطاهای پردازشی یا دستکاری داده‌ها نیز خود را نشان می‌دهد. در این موارد، از یک سو باید اعتبار داده‌ها و قابلیت استناد آن‌ها در فرایند قضایی احراز شود و از سوی دیگر، باید مشخص گردد کدام‌یک از اشخاص دخیل در طراحی، راه‌اندازی، بهره‌برداری یا نظارت بر سامانه در وقوع نتیجه زیان‌بار سهم داشته‌اند. هرچند ابزارهایی مانند بلاکچین می‌توانند در حفظ اصالت و یکپارچگی داده‌ها مؤثر باشند، اما این ابزارها به تنهایی مشکل انتساب مسئولیت را حل نمی‌کنند و همچنان نیاز به معیارهای روشن برای تعیین سهم هر عامل انسانی در ساختارهای خودکار باقی است (احمدی خونسارکی و همکاران، ۱۴۰۵، ۱۰-۹).

در چنین شرایطی، ارزیابی رکن معنوی در مواجهه با سیستم‌های تصمیم‌یار یا خودآموز به جای کشف نیت مستقیم مجرمانه، بر سه معیار عینی «امکان پیش‌بینی معقول خطر»، «اعمال احتیاط مقتضی» و «شفافیت ساختاری» متمرکز می‌شود. بر این اساس، علم و آگاهی عامل انسانی زمانی احراز می‌گردد که وی با وجود دسترسی به داده‌های خطر و توان پیش‌بینی انحرافات الگوریتم، با عدم رعایت موازین فنی و استانداردهای احتیاطی، به استقرار یا استمرار فعالیت سامانه مبادرت ورزیده باشد (ملکی، ۱۴۰۴، ۱۸۱).

۳-۲. تسبیب الگوریتمی و مسئولیت ساختاری پلتفرم‌ها

در محیط‌های دیجیتال، تحلیل رکن معنوی از جهت دیگری نیز با دگرگونی مواجه شده است و آن، نقش ساختارهای بزه‌زا در شکل‌گیری رفتار مجرمانه است. در بسیاری از موارد، رفتار زیان‌بار نتیجه یک انتخاب فردی منفک از زمینه نیست، بلکه در بستر معماری‌های فنی، سازوکارهای اقتصادی و الگوهای انگیزشی از پیش طراحی‌شده شکل می‌گیرد. در چنین فضایی، تمرکز صرف بر رفتار کاربر یا مباشر نهایی نمی‌تواند تصویر کاملی از مسئولیت کیفری به دست دهد؛ زیرا باید بررسی شود که ساختار فنی یا اقتصادی تا چه حد اشخاص را به سمت فریب، تحصیل ناروا یا دیگر رفتارهای زیان‌بار سوق داده است. این موضوع در فضای دیجیتال اهمیت بیشتری پیدا می‌کند، زیرا هویت‌های مجازی، دارایی‌های ناملموس و فرایندهای چندلایه، احراز دقیق علم، اراده و انتساب رفتار را دشوارتر می‌سازند (نماینان و هاشمپور، ۱۴۰۵، ۱۱-۱۰). از همین‌جا مفهوم «تسبیب الگوریتمی» اهمیت پیدا می‌کند. در این وضعیت، رابطه میان عامل انسانی و نتیجه مجرمانه لزوماً مستقیم و بی‌واسطه نیست، بلکه ممکن است از خلال یک سامانه خودکار، مدل تصمیم‌یار یا ساختار داده‌محور برقرار شود. هرگاه شخصی با علم به ظرفیت انحراف‌زا یا خطرآفرین سامانه، آن را طراحی، تنظیم یا به‌کارگیری کرده باشد، سببیت در قالب ارتکاب مستقیم ظاهر نمی‌شود، بلکه در قالب ایجاد و حفظ بستری تحقق می‌یابد که به‌صورت مؤثر رفتار دیگران را جهت‌دهی می‌کند. از این‌رو، تحلیل رکن معنوی در چنین مواردی باید با توجه به مفاهیم تسبیب، قاعده لاضرر و انتساب عرفی انجام گیرد (نماینان و هاشمپور، ۱۴۰۵، ۱۲-۱۰).

این سازوکار ارتباطی مستقیم با تفکیک میان «مسئولیت تقصیری» و «مسئولیت تسبیبی» دارد. در وضعیت‌هایی که مباشرت فیزیکی از ناحیه انسان منتفی است و خروجی ماشین نتیجه زیان‌بار را محقق می‌سازد، انتساب کیفری از مجرای رابطه تسبیب و قواعد فقهی نظیر قاعده اتلاف، لاضرر و ضمان محتاطانه برقرار می‌گردد. در این رویکرد، پدیدآورنده یا بهره‌بردار به دلیل ایجاد بستر پرخطر و عدم مهار علل واسطه‌ای، به عنوان سبب اقوی از مباشر، ضامن پیامدهای رفتاری سامانه شناخته می‌شود (ملکی، ۱۴۰۴، ۱۸۱-۱۸۲).

افزون بر این، تحلیل قابلیت اسناد در سامانه‌های خودآموز بدون توجه به رابطه سببیت و تقارن زمانی میان عنصر مادی و عنصر معنوی ناقص می‌ماند؛ زیرا هرچند طراحی یا آموزش سامانه در زمان پیشین صورت می‌گیرد، انتساب کیفری زمانی موجه است که حالت ذهنی لازم در زمان تحقق رفتار یا نتیجه نیز قابل احراز باشد (نحوی و همکاران، ۱۴۰۱، ۲ و ۱۱). از این‌رو، تقلیل مسئولیت به مباشر نهایی یا خروجی فنی سامانه، به انتساب ناقص و گاه گمراه‌کننده می‌انجامد و باید نقش لایه‌های پیشینی در ایجاد بستر خطرناک به‌طور مستقل ارزیابی شود (نحوی و همکاران، ۱۴۰۱، ۱-۲).

علاوه بر این، در تحلیل ساختاری تسبیب الگوریتمی، تقلیل فرایند انتساب به تصمیم‌گیرنده نهایی یا مباشر مادی، کارآمدی لازم را ندارد؛ چرا که در محیط‌های داده‌محور، جهت‌دهی نامحسوس ترجیحات و بازآرایی افق انتخاب کاربران توسط ساختار فنی صورت می‌گیرد. به همین دلیل، سنجش رکن معنوی باید به لایه‌های پیشینی شامل فرآیندهای طراحی، شیوه‌های جمع‌آوری و تنظیم داده‌ها، و نحوه اعمال معیارهای الگوریتمی تسری یابد، زیرا هرگونه تمرکز صرف بر خروجی نهایی بدون ارزیابی این بسترهای واسطه‌ای، به انتساب ناقص مسئولیت می‌انجامد (غلامی و حسینی سروری، ۱۴۰۵، ۲۷۷).

۴-۲. تقصیر ساختاری و مسئولیت توزیعی در زنجیره فناوری

گسترش فناوری هوش مصنوعی سبب شده است که مفاهیم کلاسیک حقوق کیفری، به‌ویژه در حوزه فاعلیت، انتساب و عنصر معنوی، با چالش‌های بنیادین مواجه شوند. در الگوی سنتی، مسئولیت کیفری بر مبنای وجود فاعل انسانی واجد اراده و ادراک استوار است؛ اما در بستر سامانه‌های خودآموز و تصمیم‌گیرنده، بخشی از فرایند تحقق نتیجه زیان‌بار یا مجرمانه توسط ساختارهای الگوریتمی انجام می‌شود و همین امر، الگوی متعارف انتساب را دچار ابهام می‌سازد (رضاپور، ۱۴۰۴، ۱). در این رویکرد، نتیجه

مجرمانه محصول کنش یک فاعل منفرد تلقی نمی‌شود، بلکه حاصل مجموعه‌ای از تصمیم‌ها، طراحی‌ها، سهل‌انگاری‌ها و بهره‌برداری‌های انسانی است که در بستر یک سامانه هوشمند به فعلیت می‌رسند. بنابراین، مسئولیت کیفری ممکن است میان طراح، توسعه‌دهنده، آموزش‌دهنده داده، بهره‌بردار، ناظر فنی یا شخصی که سامانه را در موقعیت پرخطر به کار گرفته است، توزیع شود.

این توزیع چندلایه مانع از آن است که بتوان تقصیر را به یک عامل انفرادی نسبت داد. از این رو، در مواجهه با زنجیره‌های پیچیده تولید فناوری، انتساب مسئولیت مستلزم گذار از مدل‌های خطی علیت به سمت «تحلیل چندعاملی» و «مسئولیت مشارکتی فناوری» است؛ رویکردی که سهم هر یک از بازیگران را در تولید، آموزش، نگهداری و نظارت بر سامانه به طور نسبی و ساختاری ارزیابی می‌کند (ملکی، ۱۴۰۴، ۱۶۵). در این بستر، پیچیدگی ساختارهای سازمانی و تعدد کنشگران در زنجیره تأمین فناوری موجب بروز پدیده‌ای تحت عنوان «مشکل دست‌های زیاد» می‌گردد؛ وضعیتی که در آن به دلیل تفکیک وظایف و توزیع مراحل تصمیم‌گیری میان طراحان، توسعه‌دهندگان و اپراتورها، احراز قصد مجرمانه یا تقصیر مستقیم فردی در یک شخص مشخص با دشواری جدی مواجه شده و ارزیابی رکن معنوی را با پیچیدگی دوچندان روبه‌رو می‌سازد (شیری و برزگر، ۱۴۰۴، ۴). در این میان، انتساب عنصر روانی به فاعل انسانی با چالش‌های نظری بنیادینی مواجه است؛ چرا که تصمیم‌گیری در سامانه‌های خودآموز بر پایه یادگیری ماشین و تحلیل مستقل کلان‌داده‌ها استوار است و خروجی رفتاری این سامانه‌ها در زمان طراحی یا بهره‌برداری، برای انسان به طور کامل پیش‌بینی‌پذیر و قابل کنترل نیست. این پیش‌بینی‌ناپذیری فنی، کارآمدی الگوهای سنتی انتساب کیفری را که بر اراده مستقیم و رابطه سببی بی‌واسطه استوارند، سلب می‌کند و ضرورت گذار به الگوهای تقصیر ساختاری را که در آن مسئولیت میان سازندگان، برنامه‌نویسان و بهره‌برداران توزیع می‌شود، آشکار می‌سازد (محمدی و رمزیار، ۱۴۰۵، ۲-۳). این جابه‌جایی مفهومی، برای بازفهم عنصر معنوی اهمیت اساسی دارد؛ زیرا سوءنیت در چنین فضایی نه لزوماً در قالب قصد مستقیم نسبت به نتیجه نهایی، بلکه گاه در قالب آگاهی به خطر، پذیرش ریسک، بی‌مبالائی حرفه‌ای یا اهمال در کنترل سامانه ظاهر می‌شود. در نتیجه، بازتعریف عنصر معنوی در بستر کنش‌های الگوریتمی مستلزم فاصله گرفتن از قرائت صرفاً انسان‌محور و خطی از رابطه فعل و نتیجه است. چنانچه شخصی با علم به ظرفیت خطرآفرین یک سامانه، آن را بدون کنترل‌های لازم طراحی، عرضه یا به کارگیری کند، زمینه انتساب کیفری منتفی نخواهد بود؛ حتی اگر بخش نهایی فرایند تصمیم‌گیری توسط خود سامانه انجام شده باشد (رضاپور، ۱۴۰۴، ۲-۳).

در این الگوی توزیعی، هرگونه بی‌توجهی به نقش خودمختار و فرآیندهای یادگیری مستقل سامانه‌های پیشرفته و اکتفا به مدل‌های سنتی انتساب، زمینه‌ساز شکل‌گیری یک خلأ مسئولیت عمیق در ساختار حقوق کیفری خواهد شد که کارآمدی واکنش‌های قانونی را به طور جدی با چالش مواجه می‌کند (فرج‌پور و همکاران، ۱۴۰۴، ۳). این خلأ ساختاری، ناشی از پدیده‌ای است که در دکترین حقوقی تحت عنوان «شکاف مسئولیت» تحلیل می‌شود؛ وضعیتی که در آن از یک سو سامانه خودمختار فاقد ادراک و اهلیت کیفری برای مجازات است و از سوی دیگر، به دلیل واسطه شدن فرایند خودتصمیم‌گیری ماشین، انتساب مستقیم جرم به تولیدکننده یا برنامه‌نویس با موانع جدی روبه‌رو می‌گردد. برای برون‌رفت از این بن‌بست نظری، مسئولیت کیفری تولیدکننده نباید صرفاً بر پایه انتساب فیزیکی فعل نهایی ربات استوار شود، بلکه باید بر مبنای ترک فعل، نقض تکالیف نظارتی و پیشگیرانه، و عدم تعبیه لایه‌های امنیتی و الگوریتم‌های مهارکننده تحلیل گردد (مستعلی، ۱۴۰۵، ۲-۳).

در این چارچوب، تقصیر ساختاری تنها ناشی از خطای فنی نیست، بلکه ریشه در معماری قدرت و نحوه بازنمایی داده‌ها دارد؛ چراکه سوگیری‌های نهفته در طراحی و انتخاب داده، مسیر تصمیم‌سازی الگوریتم را پیش از وقوع هرگونه رفتار مادی جهت‌دهی

می‌کنند. از این رو، ارزیابی مسئولیت کیفری مستلزم شفافیت در فرایند تولید دانش الگوریتمی و ردیابی دقیق زنجیره طراحی تا بهره‌برداری است تا مشخص گردد آیا ساختار فنی به گونه‌ای تنظیم شده که آگاهانه یا بر اثر غفلت، زمینه را برای وقوع نتیجه مجرمانه فراهم سازد (عباس‌زاده و پشایی، ۱۴۰۴، ۹۲-۹۳).

در این الگو، مسئولیت هر یک از بازیگران زنجیره فناوری تنها از طریق بررسی نقش آنان در تولید یا اجرای خروجی نهایی تعیین نمی‌شود، بلکه باید تکلیف آنان نسبت به کنترل، بازنگری و اصلاح تصمیمات سامانه نیز ارزیابی گردد. به ویژه، اجرای بی‌چون و چرای خروجی الگوریتم از سوی بهره‌بردار یا مقام مسئول، در حالی که امکان بررسی انسانی، اعتراض به تصمیم یا مداخله اصلاحی وجود داشته است، می‌تواند از مصادیق تقصیر نظارتی یا بی‌احتیاطی سازمانی باشد. در نتیجه، واگذاری کامل تصمیم‌گیری به سامانه، مسئولیت انسانی را زایل نمی‌کند؛ بلکه دامنه تکلیف مراقبت را به طراحی سازوکارهای نظارتی، امکان بازبینی و پاسخ‌گویی نسبت به خروجی‌های زیان‌بار گسترش می‌دهد (تنگستانی و مرادی برلیان، ۱۴۰۴، ۴۸-۵۰).

در ساختارهای نوین مبتنی بر هوش مصنوعی، تکیه بر تصمیم‌گیری‌های داده‌بنیان، ضرورت ارتقای مسئولیت‌پذیری نهادی را دوچندان کرده است. طبق رویکردهای آینده‌پژوهانه، نهادهایی که از سامانه‌های هوشمند برای تعیین تکلیف وضعیت افراد استفاده می‌کنند، نمی‌توانند صرفاً به خروجی‌های فنی بسنده کنند؛ بلکه باید حکمرانی مشارکت‌پذیر و نظارت هوشمند را به عنوان اجزای لاینفک مسئولیت‌پذیری اجتماعی خود در قبال پیامدهای این الگوریتم‌ها بپذیرند (حسینی مقدم، ۱۴۰۱، ۲ و ۱۴).

بر این اساس، شکاف مسئولیت در مواجهه با پیامدهای مجرمانه این سامانه‌ها، یک سرنوشت محتوم فناوریانه به شمار نمی‌رود، بلکه پیامد مستقیم انتخاب‌های ساختاری، طراحی‌های فنی و هنجارهای حاکم بر توسعه فناوری است. از این منظر، زمانی که سامانه‌های هوشمند بر اراده و خودمختاری تصمیم‌گیری انسان تأثیر گذاشته و آن را بازآرایی می‌کنند، مسئولیت کیفری در قالب یک عاملیت شبکه‌ای توزیع می‌شود؛ وضعیتی که در آن هرچند انسان کنترل کامل و مستقیم بر تمام مراحل پردازش را ندارد، اما به دلیل سهم خود در لایه‌های پیشین و ساختارسازی برای رفتارهای بعدی سامانه، همچنان مسئولیت حقوقی خود را حفظ می‌کند (غلامی و حسینی سروری، ۱۴۰۵، ۲۷۷).

برای برون‌رفت از این چالش حقوقی و جبران ناکارآمدی الگوهای سنتی مبتنی بر تقصیر مستقیم فردی، «دکترین مسئولیت نیابتی» به عنوان سازوکاری کارآمد برای انتساب پیامدهای رفتار هوش مصنوعی به اشخاص یا سازمان‌های به‌کارگیرنده مطرح می‌شود. این دکترین با تکیه بر مبانی نظیر ایجاد و مدیریت ریسک، ادغام سامانه در فعالیت‌های اقتصادی یا سازمانی اصیل، برتری اقتصادی اصیل برای جبران خسارت و ضرورت درونی‌سازی هزینه‌های فعالیت‌های پرخطر، امکان انتساب مسئولیت را حتی بدون نیاز به اثبات عمد یا تقصیر مستقیم فردی فراهم می‌سازد. در نظام حقوقی ایران، هرچند ماده ۱۴۲ قانون مجازات اسلامی چارچوب عمومی مسئولیت نیابتی را تبیین کرده است، اما مواجهه با ابعاد نوظهور سامانه‌های خودآموز مستلزم بازتعریف و توسعه این مفاهیم جهت پوشش جامع خلأهای مسئولیت ناشی از عاملیت الگوریتمی است (شیری و برزگر، ۱۴۰۴، ۴).

۵-۲. معیارهای هنجاری انتساب رکن معنوی به کنشگران انسانی

تفکیک میان عاملیت فنی سامانه و فاعلیت کیفری انسان ایجاب می‌کند که با عبور از نفی فاعلیت ماشین، تحلیل مسئولیت بر اشخاصی متمرکز گردد که اختیار طراحی، استقرار، کنترل یا بازنگری مؤثر بر سامانه را داشته‌اند (تنگستانی و مرادی برلیان، ۱۴۰۴، ۲۸؛ فرج‌پور و همکاران، ۱۴۰۴، ۱۹-۲۰). در این چارچوب، الگوریتم به عنوان «کنشگر فنی و واسطه معرفتی» عمل می‌کند، اما بار هنجاری رکن معنوی در طول چرخه حیات سامانه بر پایه سه معیار عینی زیر به عامل انسانی منتسب می‌گردد:

۱. کنترل مؤثر بر قلمرو ریسک و تفکیک اجزای اراده: در تحلیل رکن معنوی، تفکیک میان مؤلفه‌های قصد، اراده و آگاهی بنیادین است (قیاسی و محترم‌قلاتی، ۱۳۹۶، ۶۱۱ و ۶۳۱). اراده قوه‌ای است که رفتار خارجی را هدایت می‌کند. از این رو، در سامانه‌های خودآموز، انتساب اراده کیفری نه بر خروجی مستقیم ماشین، بلکه بر «اراده حاکم بر معماری داده، وزن‌دهی به اهداف و استقرار سامانه» استوار است. حفظ کنترل ساختاری و راهبردی مانع از انقطاع مسئولیت شده و پیوند حقوقی میان تصمیم‌گیرنده انسانی و رفتار الگوریتم را تضمین می‌کند (شیری و برزگر، ۱۴۰۴، ۱ و ۴).

۲. آگاهی هنجاری و بی‌اعتنایی آگاهانه به خطرات متعارف: آگاهی باید ناظر به واقعیت‌های مؤثر بر نتیجه زیان‌بار فهم شود. چنانچه طراح یا بهره‌بردار با وجود آگاهی از نقایص ذاتی داده‌های آموزشی، بایاس‌های سیستماتیک یا ظرفیت انحراف‌زای سامانه، آگاهانه از اصلاح یا توقف آن خودداری ورزد، این «بی‌اعتنایی به پیامدهای محتمل»، مبنای احراز رکن روانی (سوءنیت غیرمستقیم یا بی‌مبالائی مفرط) خواهد بود؛ امری که مرز میان نقص فنی محض و تقصیر کیفری را مشخص می‌سازد (نماینان و هاشمپور، ۱۴۰۵، ۱۱-۱۲).

۳. استاندارد مراقبت عینی در بستر عاملیت ترکیبی و سامانه‌های جعبه‌سیاه: در سامانه‌های پیچیده خودآموز که پیش‌بینی فنی خروجی‌های رفتاری تقلیل می‌یابد، احراز رکن معنوی بر یک «معیار عینی هنجاری» متکی است؛ بدین معنا که ملاک، قابلیت پیش‌بینی خطر و امکان مهار و بازنگری مؤثر توسط یک طراح یا بهره‌بردار متعارف و متخصص در زمان استقرار و بهره‌برداری است (مستعلی، ۱۴۰۵، ۳؛ غلامی و حسینی سروری، ۱۴۰۵، ۲۷۷).

در این فرآیند، شفافیت الگوریتمی فی‌نفسه رکن معنوی تولید نمی‌کند، بلکه صرفاً به مثابه یک واسطه اثباتی جهت سنجش «دسترسی معرفتی کنشگر» و «رعایت استاندارد مراقبت» عمل می‌نماید. بدین‌ترتیب، غیرشفاف بودن سامانه عذر موجهی برای برائت نبوده، بلکه بار نظارتی سخت‌گیرانه‌تری را برای احراز غفلت قابل انتساب بر دوش عامل انسانی بار می‌کند.

۳. بازخوانی اصول تقارن، مطابقت و تقصیر مقدم در کنش‌های واسطه‌ای و الگوریتمی

تحقق کامل مسئولیت کیفری در نظام‌های حقوقی معاصر، علاوه بر احراز ارکان مادی و معنوی، منوط به اثبات پیوند ساختاری و زمانی میان این دو رکن است. بر این اساس، صرف وجود یک رفتار زیان‌بار در عالم خارج و هم‌زمان، وجود یک ذهنیت مجرمانه یا حالت تقصیر در نفس مرتکب برای انتساب کیفری کفایت نمی‌کند، بلکه این دو عنصر باید در نقطه‌ای مشخص به یکدیگر رسیده و هم‌زمان محقق شوند؛ چرا که یک نیت یا حالت ذهنی نامطلوب تا زمانی که منشأ مستقیم رفتار خارجی قرار نگیرد، سرزنش کیفری را موجه نمی‌سازد (میرمحمدصادقی و صابری، ۱۳۹۷، ۲۲۹-۲۲۵). در تحلیل این پیوند ذهنی و مادی، سه اصل موازنه، مطابقت و تقارن زمانی نقشی تعیین‌کننده دارند. در این میان، اصل مطابقت اقتضا می‌کند که جهت‌گیری ذهنی فاعل نه تنها با ماهیت کلی جرم، بلکه در سطحی دقیق‌تر با موضوع رفتار، نتیجه حاصله و اوضاع‌واحوال حاکم بر بزه انطباق داشته باشد (میرمحمدصادقی و صابری، ۱۳۹۷، ۲۲۸).

با انتقال این مبانی به حوزه سامانه‌های خودآموز، بررسی اصل مطابقت و تقارن زمانی با پیچیدگی‌های نوینی روبه‌رو می‌شود. در یک ساختار الگوریتمی، هرگاه نتیجه مجرمانه یا زیان‌بار توسط خروجی یک سامانه یادگیری ماشین به فعلیت برسد، انتساب رکن معنوی به طراح یا کاربر منوط به آن است که قصد یا آگاهی پیشینی آن‌ها با رفتار نهایی سامانه و نتیجه حاصله مطابقت داشته باشد. به عبارت دیگر، نمی‌توان هرگونه قصد کلی برای استقرار فناوری را به طور خودکار به عنوان سوءنیت نسبت به نتایج

پیش‌بینی‌نشده یا انحرافات محاسباتی سامانه تلقی کرد، بلکه باید رابطه مستقیمی میان اراده هدایت‌کننده عامل انسانی و رفتار مادی ثبت‌شده برقرار باشد (میرمحمدصادقی و صابری، ۱۳۹۷، ۲۲۸؛ رضاپور، ۱۴۰۴، ۲).

بخش مهمی از این چالش به نحوه تفسیر «تقارن زمانی» بازمی‌گردد. در معنای سنتی، تقارن زمانی به این معناست که عنصر معنوی باید علت موجه و نیروی محرک رفتار مادی باشد و تقدم زمانی صرف، بدون وجود رابطه سببیت ذهنی، برای انتساب جرم کافی نیست (میرمحمدصادقی و صابری، ۱۳۹۷، ۲۲۶). در کنش‌های الگوریتمی، از آنجا که میان تصمیم‌سازی انسانی (در مرحله طراحی یا آموزش داده‌ها) و تحقق خروجی مجرمانه (در مرحله اجرای خودکار) فاصله زمانی و عملی وجود دارد، پیوند مستقیم میان اراده و رفتار مادی تضعیف می‌شود. در این وضعیت، برای احراز رکن معنوی باید اثبات شود که تصمیم ذهنی عامل انسانی در زمان طراحی یا راه‌اندازی، به عنوان یک نیروی محرک مستمر، در تاروپود فرایندهای محاسباتی سامانه جریان داشته و به شکلی پویا خروجی زیان‌بار را رقم زده است؛ در غیر این صورت، انتساب نتیجه به اراده اولیه به دلیل زوال تقارن و فقدان علیت ذهنی با مانع حقوقی مواجه خواهد شد (میرمحمدصادقی و صابری، ۱۳۹۷، ۲۲۷-۲۲۶؛ قیاسی و محترم‌قلاتی، ۱۳۹۶، ۶۱۱ و ۶۳۱).

با این حال، خلأ ناشی از عدم تقارن زمانی لحظه‌ای را می‌توان از طریق بازخوانی مفهوم «تقصیر مقدم» تحلیل کرد. در دکتري حقوق کیفری، هرگاه فاعل پیش از ارتکاب رفتار مجرمانه مرتکب تقصیر شده و خود را در حالتی قرار دهد که کنترل بر رفتار بعدی سلب یا محدود شود، مسئولیت کیفری او نسبت به نتایج بعدی پدیدار می‌گردد، حتی اگر در لحظه وقوع رفتار مادی، فاقد اراده مستقیم یا هوشیاری کامل باشد (میرمحمدصادقی و صابری، ۱۳۹۷، ۲۲۷). این رویکرد، بستر مناسبی را برای تحلیل مسئولیت در قبال سامانه‌های خودآموز فراهم می‌آورد. در فرضی که طراح یا کاربر با وجود آگاهی از خطرات ساختاری، نقص‌های فنی یا ظرفیت انحراف‌زای الگوریتم، آگاهانه و بدون تعبیه سوپاپ‌های امنیتی کافی، اقدام به عرضه یا فعال‌سازی سامانه در محیط‌های پرخطر می‌کند، تقصیر مقدم وی محرز است. در اینجا، هرچند در لحظه تحقق مادی جرم، مداخله انسانی مستقیمی وجود ندارد و خروجی محصول پردازش مستقل ماشین است، اما اراده مقدم بر ایجاد و استقرار مخاطره‌آمیز، پیوند معنوی لازم را برای انتساب کیفری برقرار می‌سازد (میرمحمدصادقی و صابری، ۱۳۹۷، ۲۲۷؛ نامیان و هاشمپور، ۱۴۰۵، ۱۲-۱۱).

تحلیل توزیع ضمان در سامانه‌های خودران و فناوری‌های هوشمند نشان می‌دهد که اصرار بر یافتن یک فاعل واحد و مستقیم، عملاً به ایجاد «خلأ مسئولیت» می‌انجامد. برای حل این معضل، الگوی مسئولیت لایه‌ای و سطح‌بندی‌شده پیشنهاد می‌شود که در آن، تکالیف و پاسخگویی کیفری متناسب با میزان مداخله، سطح کنترل و نقش نظارتی هر یک از کنشگران زنجیره (اعم از طراح، برنامه‌نویس و کاربر نهایی) تعریف و توزیع می‌گردد. این رژیم ترکیبی و منعطف به جای تمرکز بر فاعلیت فیزیکی محض، بر مبنای تحلیل رفتار مخاطره‌آمیز سامانه‌های خودمختار عمل کرده و پاسخگویی کیفری را در سطوح مختلف ساختار الگوریتمی مستقر می‌سازد (محمدی و رمزیار، ۱۴۰۵، ۱).

بر همین اساس، اعمال دکتري تقصیر مقدم در سامانه‌های خودآموز، با مفاهیمی چون «بی‌مبالاتی فناورانه پیشینی» و «قصور در کنترل سیستم‌های پریسک» پیوند می‌خورد. در اینجا، تقارن زمانی میان رفتار مادی و ذهنی نه در لحظه اصابت نتیجه، بلکه در زمان نقض تکالیف ایمنی، طراحی ناقص یا تأمین داده‌های انحراف‌آمیز احراز می‌شود. این بیزاری از احتیاط ساختاری در زمان طراحی یا راه‌اندازی، از منظر مبانی فقهی و حقوقی، مصداق ضمان بالمآل و تقصیر مقدم است که مسئولیت کیفری ناشی از رفتارهای خودکار بعدی ماشین را به طور کامل به کنشگران انسانی زنجیره متصل می‌سازد (ملکی، ۱۴۰۴، ۱۸۲-۱۸۱).

۴. آزمون نظریه در بستر سناریوهای عینی کنش‌های الگوریتمی

سنجش کارآمدی الگوی پیشنهادی در گرو خروج از مباحث انتزاعی و ارزیابی آن در موقعیت‌های واقعی است که در آن‌ها تصمیم‌گیری فنی سامانه با رفتارهای انسانی تلاقی می‌یابد. در این راستا، تطبیق مؤلفه‌های پنج‌گانه رکن معنوی (قصد، علم، تقصیر مقدم، کنترل مؤثر و سببیت هنجاری) در سه حوزه متمایز، نحوه انتساب هنجاری را آشکار می‌سازد:

۱-۴. سامانه‌های خودران و سوانح ترافیکی مرگبار

در فرض وقوع حادثه مرگبار توسط یک خودروی خودران (سطح چهار یا پنج) ناشی از عدم تشخیص عابر پیاده در شرایط خاص نوری یا جوی، مواجهه با رکن معنوی ماهیتی چندلایه می‌یابد. در این سناریو، قصد مستقیم ارتکاب جنایت از سوی تیم مهندسی یا تولیدکننده مفقود است؛ لکن احراز رکن معنوی بر پیوند میان «علم پیشینی به ریسک» و «تقصیر مقدم در فرآیند استقرار» استوار می‌گردد. چنانچه مستندات فنی و آزمایشی فاز توسعه دلالت بر این داشته باشد که سامانه در مواجهه با شرایط جوی نامتعارف دارای خطای ادراکی سیستماتیک بوده و تولیدکننده با وجود آگاهی از این نقیصه، به منظور رقابت تجاری و بدون تعبیه لایه‌های نظارتی بازدارنده اقدام به عرضه سامانه نموده باشد، رفتار وی مصداق بی‌مبالاتی مفرط و ایجاد مخاطره نامتعارف خواهد بود. از منظر سببیت هنجاری، استقلال عملیاتی خودرو در لحظه وقوع تصادف مانع از انتساب نتیجه به شرکت سازنده نیست؛ زیرا کنترل مؤثر نه در زمان وقوع حادثه، بلکه در مرحله معماری داده‌ها و مدیریت ریسک اعمال می‌شده است. بدین ترتیب، خروجی الگوریتم صرفاً مجرای تحقق خطری است که منشأ آن به اراده و قصور انسانی در مراحل پیشینی بازمی‌گردد و بر مبنای قاعده تسبیب، زمینه انتساب جنایت شبه‌عمدی یا خطای محض ناشی از بی‌احتیاطی را فراهم می‌آورد.

۲-۴. سامانه‌های تصمیم‌یار پزشکی و خطاهای تشخیصی-درمانی

در حوزه پزشکی، کاربرد الگوریتم‌های یادگیرنده در تحلیل تصاویر رادیولوژی یا پیشنهاد پروتکل‌های درمانی پیچیده، مسئله تفکیک مرز میان «خطای فنی ابزار» و «تقصیر جزایی پزشک» را مطرح می‌سازد. در سناریویی که یک سامانه تصمیم‌یار دوز دارویی مرگباری را به دلیل انحراف در تحلیل تداخلات به بیمار پیشنهاد می‌دهد و پزشک معالج بدون بازبینی نقادانه و تطبیق با موازین بالینی آن را تأیید می‌کند، رکن معنوی مستقیماً بر محور مفهوم «سوگیری ناشی از اتکای خودکار» و نقض تکلیف مراقبت تحلیل می‌شود.

در این فرض، اگرچه قصد سوء در رفتار پزشک وجود ندارد، اما آگاهی هنجاری وی از جایگاه مشورتی سامانه و لزوم استقلال در تصمیم‌گیری بالینی، موجب شکل‌گیری تکلیف قانونی نسبت به راستی‌آزمایی خروجی می‌شود. پزشک با در اختیار داشتن حق و امکان بازنگری و لغو دستور سامانه، واجد کنترل مؤثر بوده و تسلیم غیرانتقادی وی در برابر پیشنهاد الگوریتم، مصداق بارز بی‌احتیاطی ناشی از غفلت است. از این رو، به دلیل مداخله هنجاری ناقص پزشک در زنجیره تصمیم‌گیری، رابطه سببیت به رفتار مباشرتی وی متصل گردیده و مسئولیت کیفری ناشی از فوت بیمار بر مبنای خطای تشخیصی غیرقابل اغماض به وی منتسب خواهد شد؛ مگر آنکه نقص سامانه ناشی از پنهان‌کاری فنی طراح در شفاف‌سازی محدودیت‌های مدل باشد که در این حالت، مسئولیت در قالب تسبیب طولی میان طراح و پزشک قابل توزیع است.

۳-۴. معاملات الگوریتمی خودکار و اخلال در سامانه‌های اقتصادی

در سامانه‌های مبادلاتی با بسامد بالا یا پلتفرم‌های توزیع کالا، پیاده‌سازی الگوریتم‌های یادگیرنده‌ای که با هدف حداکثرسازی سود طراحی شده‌اند، ممکن است به شکل‌گیری الگوهای رفتاری انحصاری، ارسال سفارش‌های صوری و دستکاری غیرقانونی قیمت‌ها بینجامد. در این موقعیت، تحلیل رکن معنوی مدیران و طراحان به سوی احراز «قصد احتمالی یا تبعی» سوق می‌یابد.

هنگامی که بهره‌برداران با علم به رفتار تهاجمی الگوریتم و استعداد بالای آن در تخطی از قواعد رقابتی، تابع پاداش سامانه را به صورت لجام‌گسیخته تعریف کرده و هیچ‌گونه سازوکار مهار خودکار یا توقف اضطراری برای جلوگیری از رفتارهای غیرقانونی تعبیه نکنند، استناد به جعبه‌سیاه بودن یا غیرقابل پیش‌بینی بودن ماشین مسموع نخواهد بود. در این بستر، عدم شفافیت فنی نه تنها رافع مسئولیت نیست، بلکه پذیرش آگاهانه احتمال وقوع جرم و بی‌اعتنایی نسبت به صیانت از نظم بازار، نشان‌دهنده اراده ضمنی در تحقق نتیجه مجرمانه است. از این رو، اخلال در بازار یا کلاه‌برداری رایانه‌ای از طریق تسبیب معنوی به مدیرانی منتسب می‌گردد که با تعریف اهداف مخاطره‌آمیز، ابزار فنی را به سوی رفتارهای انحرافی سوق داده‌اند.

نتیجه‌گیری

بررسی تحلیلی مسئله نشان می‌دهد که ورود کنش‌های الگوریتمی و سامانه‌های خودآموز، نه به زوال مسئولیت کیفری می‌انجامد و نه امکان تداوم بی‌واسطه الگوی کلاسیک انتساب را باقی می‌گذارد؛ بلکه ضرورت بازسازی معیارهای احراز رکن معنوی را در سطحی فراتر از تقابل ساده «فاعل انسانی/ابزار فنی» آشکار می‌کند. در این چارچوب، سامانه هوشمند هرچند می‌تواند در تولید، تسریع یا تشدید نتیجه مجرمانه نقش علی و فنی داشته باشد، اما به دلیل فقدان اراده هنجاری، ادراک اخلاقی و قابلیت سرزنش، فاعل کیفری مستقل محسوب نمی‌شود (انصاری و مهدیان، ۱۴۰۴، ۲۲۸؛ ملکی، ۱۴۰۴، ۱۵۹-۱۶۰). از این رو، محور تحلیل مسئولیت باید از خود خروجی الگوریتم به رفتار پیشینی اشخاص و نهادهای انسانی منتقل شود؛ اشخاصی که در مقام طراح، برنامه‌نویس، تأمین‌کننده داده، بهره‌بردار، ناظر یا تصمیم‌گیرنده، در شکل‌دهی به معماری خطر نقش داشته‌اند (ملکی، ۱۴۰۴، ۱۸۱-۱۸۲).

یافته مهم دیگر آن است که عنصر روانی در این حوزه، الزاماً به قصد مستقیم یا علم قطعی به نتیجه محدود نمی‌شود؛ بلکه می‌تواند در قالب علم به خطر قابل پیش‌بینی، پذیرش آگاهانه ریسک، بی‌احتیاطی ساختاری، غفلت مهندسی، ترک نظارت مؤثر و استمرار بهره‌برداری از سامانه‌ای با نقص شناخته‌شده متجلی شود (ملکی، ۱۴۰۴، ۱۶۵؛ تنگستانی و مرادی برلیان، ۱۴۰۴، ۲۲ و ۴۹). بر این اساس، رکن معنوی جرم در بستر فناوری‌های خودآموز باید با مفاهیمی چون تقصیر مقدم، تسبیب الگوریتمی، مسئولیت توزیعی و علیت چندلایه بازفهم شود؛ مفاهیمی که امکان انتساب منصفانه را در زنجیره پیچیده تولید و بهره‌برداری از سامانه‌های هوشمند فراهم می‌کنند (ملکی، ۱۴۰۴، ۱۵۷-۱۵۹).

در نتیجه، بازتعریف رکن معنوی در این حوزه، به معنای توسعه معیارهای انتساب کیفری بر پایه کنترل مؤثر، قابلیت پیش‌بینی، امکان مداخله، شفافیت و نظارت انسانی است. هر جا این عناصر در سطح لازم وجود نداشته باشند، نه تنها احراز تقصیر انسانی دشوارتر می‌شود، بلکه خطر شکل‌گیری «شکاف مسئولیت» نیز افزایش می‌یابد. بنابراین، تحلیل کیفری کنش‌های الگوریتمی باید از منطق انتساب خطی فاصله گرفته و به الگوی مسئولیت شبکه‌ای و چندعاملی حرکت کند؛ الگویی که هم از سرزنش بی‌پایه سامانه جلوگیری می‌کند و هم مانع از گریز کنشگران انسانی از مسئولیت در پس پیچیدگی فناوریانه می‌شود.

پیشنهادهای

با توجه به پیچیدگی انتساب کیفری در بستر سامانه‌های خودآموز، تدوین چارچوب قانونی ویژه برای تنظیم مسئولیت ناشی از کنش‌های الگوریتمی ضروری است. این چارچوب باید ضمن حفظ اصل شخصی بودن مسئولیت کیفری و منع مسئولیت عینی، معیارهای روشن احراز نقش انسانی در طراحی، آموزش، استقرار، بهره‌برداری و نظارت بر سامانه را تعیین کند. در این زمینه، صرف وقوع نتیجه زیان‌بار نباید مبنای انتساب مسئولیت قرار گیرد؛ بلکه باید رابطه میان نقش شخص، کنترل مؤثر، قابلیت پیش‌بینی خطر، امکان مداخله و نقض تکلیف احتیاطی احراز شود.

نخست، پیشنهاد می‌شود تکالیف پیشینی و حرفه‌ای اشخاص و نهادهای دخیل در چرخه حیات سامانه‌های هوشمند به‌طور صریح تعیین شود. ارزیابی خطر پیش از استقرار، اعتبارسنجی داده‌های آموزشی، آزمون عملکرد در موقعیت‌های پرخطر، ثبت تغییرات مدل، مستندسازی تصمیم‌های فنی، پایش مستمر خروجی‌ها و پیش‌بینی سازوکار توقف اضطراری، از جمله تکالیفی است که نقض آن‌ها می‌تواند در احراز تقصیر یا تسبیب کیفری مؤثر باشد. هرچه احتمال ورود آسیب به حقوق بنیادین، جان، مال، آزادی یا حریم خصوصی اشخاص بیشتر باشد، سطح تکلیف به مراقبت، شفافیت و نظارت نیز باید افزایش یابد (تنگستانی و مرادی برلیان، ۱۴۰۴، ۳۴ و ۴۹).

دوم، لازم است نظام «ردیابی‌پذیری و قابلیت ممیزی الگوریتمی» در حوزه‌های حساس پیش‌بینی شود. نگهداری داده‌های مربوط به منبع و کیفیت داده‌های آموزشی، نسخه‌های مدل، تغییرات اعمال‌شده، هشدارهای صادرشده، تصمیم‌های اپراتور انسانی و رخدادهای زمان اجرای سامانه، امکان بازسازی مسیر علی‌منتهی به نتیجه زیان‌بار را فراهم می‌کند. بدون چنین سازوکاری، احراز رابطه سببیت، تعیین قلمرو کنترل اشخاص و تشخیص علم یا بی‌احتیاطی آنان با دشواری جدی مواجه خواهد شد.

سوم، پیشنهاد می‌شود در تنظیم مسئولیت کیفری، میان نقش‌های متفاوت بازیگران زنجیره فناوری تفکیک شود. مسئولیت طراح یا توسعه‌دهنده، تأمین‌کننده داده، عرضه‌کننده، مالک، بهره‌بردار و ناظر، تابع سطح اختیار، آگاهی، کنترل و تکلیف قانونی هر یک است و نباید صرفاً بر مبنای عنوان یا موقعیت سازمانی آنان تعیین شود. تحلیل چندعاملی و مسئولیت توزیعی، جایگزین مناسب‌تری برای انتساب خطی نتیجه به یک شخص واحد است؛ زیرا تولید خطر در سامانه‌های خودآموز غالباً محصول تصمیم‌ها و ترک‌فعل‌های متراکم و چندمرحله‌ای است (ملکی، ۱۴۰۴، ۱۸۱-۱۸۲).

چهارم، ضروری است مفهوم تقصیر در حقوق کیفری با اقتضائات فنی سامانه‌های خودآموز متناسب‌سازی شود. بی‌توجهی به هشدارهای فنی، استفاده از داده‌های معیوب یا سوگیرانه، استقرار سامانه بدون آزمون کافی، حذف سازوکار نظارت انسانی، ادامه بهره‌برداری پس از آشکارشدن نقص و امتناع از اصلاح مدل، باید حسب مورد از مصادیق بی‌احتیاطی، بی‌مبالاتی یا تقصیر مقدم تلقی شود. این رویکرد، بدون پذیرش شخصیت کیفری برای سامانه، امکان پاسخ‌گویی نسبت به خطرهایی را فراهم می‌آورد که در مراحل پیش از وقوع نتیجه ایجاد یا تشدید شده‌اند (ملکی، ۱۴۰۴، ۱۸۱-۱۸۲).

پنجم، پیشنهاد می‌شود در حوزه‌های دارای خطر بالا، مانند حمل‌ونقل خودران، خدمات درمانی، تصمیم‌گیری‌های مالی، زیرساخت‌های حیاتی و سامانه‌های مؤثر بر حقوق و آزادی‌های اشخاص، اصل «نظارت انسانی مؤثر» به‌عنوان یک الزام قانونی پذیرفته شود. نظارت انسانی نباید صوری، غیرقابل اجرا یا فاقد اختیار مداخله باشد؛ بلکه ناظر باید به اطلاعات کافی، امکان توقف یا اصلاح تصمیم و صلاحیت فنی لازم دسترسی داشته باشد. در غیر این صورت، حضور ظاهری انسان نمی‌تواند مانع از تحقق تقصیر ساختاری شود.

در نهایت، توسعه پژوهش‌های میان‌رشته‌ای میان حقوق کیفری، علوم رایانه، مهندسی داده، اخلاق فناوری و فقه اسلامی ضرورت دارد. ظرفیت قواعدی چون تسبیب، لاضرر و ضمان ناشی از ایجاد خطر، می‌تواند در تبیین مسئولیت ناشی از طراحی ناقص، ترک نظارت و ایجاد وضعیت خطرناک نقش مؤثری ایفا کند؛ مشروط بر آن‌که کاربرد این مبانی با اصول قانونی بودن جرم و مجازات، شخصی بودن مسئولیت و لزوم احراز تقصیر سازگار شود (ملکی، ۱۴۰۴، ۱۸۱-۱۸۲).

منابع:

- (۱) احمدی خونسارکی، پریسا، عباسی، معصومه، و کشکولیان، اسماعیل. (۱۴۰۵). تأثیر فناوری‌های نوین در کشف و پیشگیری جرایم زیست‌محیطی. فصلنامه علمی فقه و حقوق نوین، ۷(۲۶)، ۱-۲۲.

- (۲) انصاری، منصور، و مهدانیا، سامان. (۱۴۰۴). هوش مصنوعی و تحول سوژکتیویته سیاسی: در جست‌وجوی یک راه‌حل. دوفصلنامه علمی پژوهش سیاست نظری، (۳۷)، ۲۲۳-۲۵۲.
- (۳) تنگستانی، محمد قاسم، و مرادی برلیان، مهدی. (۱۴۰۴). مطالعه تطبیقی مشروعیت حقوقی تصمیمات الگوریتمی در دولت‌های دیجیتال با تمرکز بر تجربه حقوقی ایران. فصلنامه حقوق اداری، ۱۳(۴۴)، ۲۷-۵۲.
- (۴) رضاپور، جعفر. (۱۴۰۴). بررسی تطبیقی واکنش کیفری به جرایم ناشی از هوش مصنوعی در ایران و اتحادیه اروپا. نشریه علمی پژوهشی پژوهشگران بین‌المللی حقوق، (۵)، ۱-۲۰.
- (۵) شهسواری، احسان. (۱۴۰۲). مطالعه تطبیقی الگوی ایرانی «حمایت از کودکان و نوجوانان در فضای مجازی»: ارزیابی از منظر استانداردهای بین‌المللی حقوق کودک. نشریه حقوق تطبیقی، ۷(۲)، ۱۰۵-۱۲۴.
- (۶) شیر، عباس، و برزگر، محمدرضا. (۱۴۰۴). تحلیل دکرین مسئولیت نیابتی به مثابه مبنای مسئولیت ناشی از هوش مصنوعی در حقوق ایران، با الهام از رویه قضایی کامن‌لا. فصلنامه حقوق، ۵۵ (پذیرفته‌شده)، ۱-۲۸.
- (۷) عباس‌زاده، محمد، و پاشایی، سجاد. (۱۴۰۴). مصاحبه‌گران الگوریتمی و بازاندیشی در تولید دانش: به‌سوی روش‌شناسی انتقادی هوشمند در پژوهش‌های کیفی مبتنی بر هوش مصنوعی. روش‌شناسی علوم انسانی، ۳۱(۱۲۳)، ۸۳-۱۰۵.
- (۸) فرج‌پور، رضا، محمودی‌مقدم، سید مصطفی، بابایی، داریوش، رستمی، حمیدرضا، و کیانی، محمدجواد. (۱۴۰۴). آثار و پیامدهای حقوقی اعطای شخصیت حقوقی به سیستم‌های هوش مصنوعی در حقوق ایران و اتحادیه اروپا. پژوهش‌های نوین در مطالعات علوم انسانی اسلامی، ۴(ویژه‌نامه)، ۱-۲۸.
- (۹) قربانی‌درزی‌محله، افسانه، و سبحانی، مهین. (۱۴۰۳). بررسی ماهیت و انواع فاعل معنوی در دیوان کیفری بین‌المللی. مجله حقوقی بین‌المللی، ۴۱(۷۴)، ۱۰۵-۱۳۲.
- (۱۰) قیاسی، جلال‌الدین، و محترم‌قلاتی، ایمان. (۱۳۹۶). تحلیل عنصرمحور رکن معنوی جنایات عمدی در صلاحیت دیوان کیفری بین‌المللی. فصلنامه پژوهش حقوق کیفری، ۶(۲۱)، ۶۱۱-۶۳۱.
- (۱۱) غلامی، طیب، و حسینی سروری، سید حسن. (۱۴۰۵). هوش مصنوعی و دگرگونی خودمختاری انسان: تحلیلی از اراده آزاد و مسئولیت اخلاقی در تصمیم‌سازی داده‌محور. پژوهش‌های فلسفی، ۲۰(۵۵)، ۲۶۵-۲۸۴.
- (۱۲) ملکی، محمدمبین. (۱۴۰۴). تحلیل فقهی - حقوقی کیفرپذیری رفتارهای ناشی از «خطای الگوریتمی» در جرایم دیجیتال: از تقصیر تا مسئولیت تسبیبی. مجلس و قانونگذاری، ۱(۲)، ۱۵۷-۱۸۴.
- (۱۳) محمد حسینی مقدم. (۱۴۰۱). روندهای جهانی علم و فناوری، راهبردها و اولویت‌های آینده‌پژوهی آموزش عالی در ایران. دوفصلنامه علمی آینده‌پژوهی ایران، ۶(۲)، ۱-۲۸.
- (۱۴) محمدی ورمزیار، سلمان. (۱۴۰۵). تحلیل مسئولیت کیفری در حوزه هوش مصنوعی و سامانه‌های خودران. ماهنامه علمی تخصصی پایا شهر، (۸۷)، ۱-۲۰.
- (۱۵) مستعلی، سمانه. (۱۴۰۵). مسئولیت کیفری تولیدکنندگان ربات‌های خودمختار در جرایم بدون عامل انسانی مستقیم. مطالعات علوم سیاسی، حقوق و فقه، ۱۲(۱)، ۵۳۷-۵۴۸.
- (۱۶) میرمحمدصادقی، حسین، و صابری، راضیه. (۱۳۹۷). تقارن زمانی عناصر مادی و معنوی جرم: ماهیت و رویکردها. فصلنامه دیدگاه‌های حقوق قضایی، ۲۳(۸۴)، ۲۲۵-۲۶۲.

- (۱۷) نحوی، عادل، قربانی، علی، و عباسی، اصغر. (۱۴۰۱). ارکان تشکیل‌دهنده جرم و قابلیت اسناد در حقوق کیفری انگلستان. فصلنامه فقه جزای تطبیقی، ۲(۵)، ۱۴۵-۱۵۶.
- (۱۸) نمامیان، پیمان، و هاشمپور، مدیحه. (۱۴۰۵). چالش‌های فقه اقتصادی در مواجهه با بزهکاری‌های مالی بازی‌های رایانه‌ای برخط. فصلنامه فقه جزای تطبیقی، ۶(۱)، ۱-۱۵.