

بررسی چالش های حقوق کیفری در مواجهه با جرائم مبتنی بر هوش مصنوعی

طلا لیراوی^۱*

کارشناس ارشد حقوق جزا و جرم شناسی، دانشگاه آزاد بوشهر. (کارشناس حقوقی شهرداری).

چکیده

گسترش سریع فناوری های مبتنی بر هوش مصنوعی در دهه های اخیر، ابعاد مختلف زندگی اجتماعی، اقتصادی و حقوقی را تحت تأثیر قرار داده و در کنار مزایای گسترده، چالش های تازه ای نیز برای نظام های حقوقی به ویژه حقوق کیفری ایجاد کرده است. استفاده از سامانه های هوشمند در حوزه هایی مانند تولید محتوا، تصمیم گیری خودکار، تحلیل داده های کلان و تعاملات دیجیتال، زمینه بروز گونه های جدیدی از رفتارهای مجرمانه را فراهم کرده که تحلیل و پاسخ دهی به آن ها در چارچوب مفاهیم سنتی حقوق کیفری با دشواری هایی همراه است. از این رو، بررسی چالش های حقوق کیفری در مواجهه با جرائم مبتنی بر هوش مصنوعی به یکی از موضوعات مهم در مطالعات حقوقی معاصر تبدیل شده است. هدف اصلی این پژوهش، تبیین مهم ترین چالش های حقوق کیفری در مواجهه با جرائم مبتنی بر هوش مصنوعی و تحلیل رویکردهای حقوقی موجود برای مدیریت این پدیده نوظهور است. روش پژوهش حاضر از نوع توصیفی-تحلیلی بوده و با بهره گیری از روش مطالعه کتابخانه ای انجام شده است. در این راستا، مقالات علمی، منابع حقوقی و مطالعات پژوهشی مرتبط با موضوع هوش مصنوعی و حقوق کیفری مورد بررسی و تحلیل قرار گرفته و تلاش شده است با رویکردی تحلیلی، مهم ترین مسائل نظری و عملی در این حوزه شناسایی و طبقه بندی شود. همچنین با استفاده از رویکرد تطبیقی، برخی سیاست ها و رویکردهای اتخاذ شده در نظام های حقوقی مختلف مورد توجه قرار گرفته است تا تصویر روشن تری از شیوه های مواجهه حقوق کیفری با این نوع جرائم ارائه شود. یافته های پژوهش نشان می دهد که جرائم مبتنی بر هوش مصنوعی چالش های متعددی را در حوزه هایی همچون تعیین فاعل جرم، احراز عنصر روانی، اثبات رابطه علیت و فرآیند کشف و اثبات جرم ایجاد کرده اند. پیچیدگی الگوریتم ها، ماهیت خودیادگیر برخی سامانه های هوشمند و ساختار چندعاملی فرآیندهای فناورانه، موجب شده است که انتساب رفتار مجرمانه به یک شخص مشخص با دشواری های جدی مواجه شود. همچنین بررسی رویکردهای تطبیقی نشان می دهد که بسیاری از نظام های حقوقی در تلاش اند با ترکیبی از اصلاحات تقنینی، تنظیم گری پیشگیرانه و تقویت نظارت بر فناوری، چارچوبی مناسب برای مدیریت مخاطرات هوش مصنوعی ایجاد کنند.

واژگان کلیدی: هوش مصنوعی، جرائم سایبری، مسئولیت کیفری، چالش های حقوقی، سیاست جنایی

۱. مقدمه

تحولات شتابان فناوری های نوین، به ویژه در حوزه هوش مصنوعی، الگوهای سنتی ارتکاب جرم، کشف آن و واکنش کیفری را به طور بنیادین دگرگون ساخته و ذهنیت کلاسیک حقوق دان کیفری را با پرسش ها و چالش های کم سابقه ای روبه رو کرده است. هوش مصنوعی دیگر صرفاً یک ابزار کمک کننده در دست انسان نیست، بلکه در قالب سامانه ها و الگوریتم های خودیادگیر، شبکه های عصبی عمیق، ربات های خودمختار و خودروهای خودران، در تصمیم گیری های حساس، پرخطر و بعضاً غیرقابل پیش بینی مداخله می کند و بدین ترتیب، نسبت میان فعل انسان، رفتار سامانه و نتیجه زیان بار، دچار پیچیدگی های بی سابقه می شود (میرلوحی و همکاران، ۲۰۲۵). در چنین بستری، پرسش محوری این است که حقوق کیفری، که بر مبنای مفاهیم سنتی «شخص»، «اراده»، «قصد» و «تقصیر» بنا شده، تا چه حد توان و ظرفیت مواجهه با جرائم مبتنی بر هوش مصنوعی را دارد و آیا ابزارهای موجود آن، از جمله عناصر سه گانه جرم، قواعد انتساب نتیجه، ساختار مسئولیت کیفری و سازوکارهای اثبات، برای تنظیم و کنترل این واقعیت نوظهور کفایت می کند یا خیر (زندى و همکاران، ۲۰۲۴). در حقیقت، ورود هوش مصنوعی به عرصه های متنوعی چون حمل و نقل خودران، خدمات مالی و بانکی، تشخیص و درمان پزشکی، نظامی و امنیتی، پلیس هوشمند، نظارت فراگیر و حتی فرایند دادرسی و قضاوت، موجب شده است که جرائم بالقوه، نه فقط در سطح جرائم سایبری متعارف، بلکه در قالب جرائم سازمان یافته، تروریستی، اقتصادی و حتی بین المللی متجلی گردد و از این رهگذر، مرز میان «فاعل انسانی» و «سیستم هوشمند» در ارتکاب رفتار مجرمانه تیره و مبهم شود (خلیلی پاچی و همکاران، ۲۰۲۵).

اهمیت و ضرورت بررسی چالش های حقوق کیفری در مواجهه با جرائم مبتنی بر هوش مصنوعی، از چند جهت قابل تبیین است. نخست آن که توسعه روزافزون سامانه های هوشمند و حضور آن ها در فرآیندهای زیربنایی و حیاتی، مانند زیرساخت های انرژی، حمل و نقل عمومی، شبکه های مالی و ارتباطی، به گونه ای است که هرگونه اختلال، سوءاستفاده یا خطای سیستمی می تواند نتایج ویرانگر گسترده ای برای امنیت فردی و جمعی به بار آورد و در عین حال، همواره پرسش از اینکه «چه کسی باید پاسخ گو باشد؟» بی پاسخ یا دست کم مبهم باقی می ماند (کریمی، ۲۰۲۵). دوم آن که، بسیاری از نظام های حقوقی، از جمله ایران، هنوز فاقد یک چارچوب تقنینی جامع و صریح در خصوص تنظیم حقوقی هوش مصنوعی و مسئولیت ناشی از کاربرد آن هستند و به ناچار، برای پاسخ گویی به مسائل نو، به قواعد عام و سنتی حقوق کیفری و همچنین قیاس و توسعه مفهومی متوسل می شوند؛ امری که می تواند با اصول بنیادین حقوق کیفری مانند قانونی بودن جرائم و مجازات ها، تفسیر مضیق قوانین کیفری و منع قیاس در تعارض قرار گیرد (رضوی فرد و همکاران، ۲۰۱۶). سوم آن که، از منظر سیاست جنایی، بی توجهی به این تحولات می تواند موجب تعمیق «شکاف تنظیمی» گردد؛ شکافی که در آن، کارگزاران فناوری و مرتکبان بالقوه، از خلأ یا ابهام قانونی بهره برداری نموده و متضرران و جامعه، در فقدان سازوکارهای روشن جبران و بازدارندگی، در وضعیت بی دفاع باقی می مانند (سردارنیا و همکاران، ۲۰۱۹).

در سطح نظری، مسئله هوش مصنوعی، بنیان های فلسفی و مفهومی مسئولیت کیفری را نیز به چالش می کشد. نظام های حقوق کیفری کلاسیک، بر فرض «شخص بودن فاعل» و برخوردارى او از اراده و شعور استوار است؛ به نحوی که امکان نسبت دادن قصد یا تقصیر و در پی آن، سرزنش اخلاقی و حقوقی فراهم شود. حال آن که در مورد سامانه های هوش مصنوعی پیشرفته، خصوصاً سامانه های خودیادگیر و خودمختار، پرسش این است که آیا می توان به طور مستقیم به سامانه، «اراده» یا «قصد» نسبت داد یا آن که همچنان باید در پی یافتن شخص حقیقی یا حقوقی پشت سامانه بود و مسئولیت را به طراح، برنامه نویس، تولید کننده، بهره بردار، ناظر یا ترکیبی از آن ها نسبت داد (زندى و همکاران، ۲۰۲۴). برخی دیدگاه ها، با الهام از نظریه های مسئولیت مبتنی

بر ریسک و نیز تحلیل های نوین در فقه امامیه و دیگر مکاتب فقهی، بر این باورند که می توان از مسئولیت صرفاً شخص محور به سوی مسئولیت مبتنی بر ریسک و تخصیص خطر گذار کرد؛ به این معنا که هرکس با بهره گیری از سامانه های هوشمند، سطحی از خطر را وارد فضای اجتماعی می کند، باید متناسب با آن خطر، تعهدات و مسئولیت های کیفری و شبه کیفری خاصی را بر عهده گیرد، حتی اگر رابطه مستقیم و سنتی میان فعل و قصد او با نتیجه زیان بار احراز نشود (میرلوحی و همکاران، ۲۰۲۵؛ صمدیان و همکاران، ۲۰۲۵). چنین رویکردی، البته با پرسش هایی در باب سازگاری با مبانی فقهی و اصول حاکم بر حقوق کیفری ایران روبه رو است و نیازمند واکاوی دقیق استدلالی و تطبیقی است (صمدیان و همکاران، ۲۰۲۵).

هم زمان، تجربه نظام های حقوقی پیشرو در مواجهه با فناوری های مبتنی بر هوش مصنوعی، از جمله در حوزه خودروهای خودران و وسایل نقلیه خودکار، نشان می دهد که حتی در نظام های حقوقی توسعه یافته نیز، مسئله انتساب مسئولیت کیفری و تنظیم رفتار این سامانه ها، محل اختلاف و چالش های جدی است. برای مثال، در آلمان و برخی کشورهای اروپایی، تلاش هایی برای وضع چارچوب های تقنینی ویژه خودروهای خودران صورت گرفته که در آن ها، به نحوی، توزیع مسئولیت میان سازنده، مالک، بهره بردار و حتی نهادهای نظارتی مدنظر قرار گرفته است، اما باز هم پرسش از کفایت این مقررات در مواجهه با سناریوهای پیچیده و پیش بینی ناپذیر، همچنان مطرح است (خوشان تاش و همکاران، ۲۰۲۵). مطالعات تطبیقی در خصوص مقررات کیفری مرتبط با هوش مصنوعی و وسایل نقلیه خودران در نظام های حقوقی سنگاپور، فرانسه و ایران نیز نشان می دهد که علی رغم برخی تلاش ها برای تنظیم این حوزه، هنوز هم در زمینه هایی چون معیارهای تقصیر، استانداردهای ایمنی، الزامات شفافیت الگوریتمی، تکالیف گزارش دهی و نظام های بیمه ای و جبرانی، خلأها و ابهامات مهمی وجود دارد (الذبحاوی و همکاران، ۲۰۲۵). این وضعیت، برای قانون گذار ایرانی، هم هشداردهنده و هم فرصت ساز است؛ هشدار از جهت عقب ماندن در تنظیم حقوقی پدیده های که به سرعت در حال تعمیم به عرصه های مختلف زندگی است و فرصت از جهت امکان بهره گیری از تجربه دیگر کشورها و تطبیق آن با مبانی حقوقی و فقهی بومی (خوشان تاش و همکاران، ۲۰۲۵). از سوی دیگر، جرائم مبتنی بر هوش مصنوعی تنها به حوزه حمل و نقل یا خطاهای فنی محدود نمی شود، بلکه در حوزه های حساس امنیتی، مالی و تروریستی نیز بروز یافته است. استفاده از هوش مصنوعی در طراحی و اجرای حملات سایبری پیچیده، عملیات اطلاعاتی و روانی، هدایت پهپادها و سامانه های تسلیحاتی خودکار، و نیز تسهیل جریان های مالی غیرقانونی و پول شویی در بستر رمزارزها و پلتفرم های غیرمتمرکز، نمونه هایی از این ظرفیت مجرمانه اند (عبدالهی و همکاران، ۲۰۲۵). در چنین فضایی، هم گروه های تروریستی و هم شبکه های جنایت کار سازمان یافته، می توانند از هوش مصنوعی برای افزایش دقت، پوشش و ناشناسی عملیات خود استفاده کرده و کارآمدی سازوکارهای سنتی کشف و پیشگیری را به چالش بکشند (عبدالهی و همکاران، ۲۰۲۵؛ کاهدی و همکاران، ۲۰۲۵). برای مثال، کاربرد الگوریتم های یادگیری ماشین در شست و شوی وجوه ناشی از جرائم سازمان یافته، از طریق ساخت الگوهای پیچیده تراکنشی در رمزارزها، نه تنها ردیابی منشأ وجوه را دشوار می کند، بلکه می تواند قوانین موجود در حوزه مبارزه با پول شویی را که بر شناسایی الگوهای تکراری و غیرعادی استوار است، ناکارآمد سازد (کاهدی و همکاران، ۲۰۲۵). در سطح خردتر، نیز استفاده از هوش مصنوعی در نقض حریم خصوصی، تولید محتوای عمیق جعلی، دست کاری هویت ها در متاورس و فضاهای مجازی مشابه، زمینه ساز گونه های جدیدی از تعرض به حقوق فردی و کرامت انسانی است که نظام کیفری باید برای آن ها پاسخ گو باشد (کریمی، ۲۰۲۵).

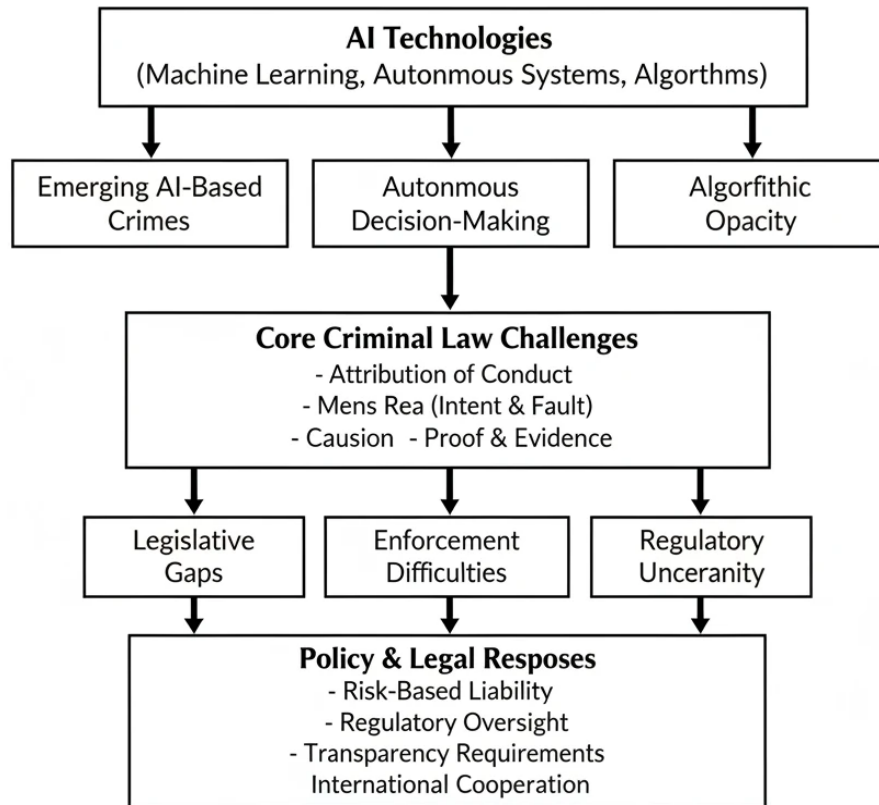
با توجه به این تحولات، پرسش های اساسی متعددی پیش روی حقوق دان کیفری قرار می گیرد: نخست آن که، در مواجهه با جرائمی که مستقیماً از رفتار یا تصمیم یک سامانه هوش مصنوعی ناشی می شود، چگونه می توان رابطه علیت و انتساب کیفری

را میان دستور یا طراحی اولیه انسان و نتیجه زیان بار برقرار کرد؟ آیا صرف طراحی یک الگوریتم، بدون پیش بینی خاص از نحوه رفتار آن در موقعیت های آینده، برای انتساب نتیجه کفایت می کند یا باید معیاری از پیش بینی پذیری عینی یا نوعی را ملاک قرار داد (میرلوحی و همکاران، ۲۰۲۵)؟ دوم آن که، در قلمرو عنصر معنوی، آیا می توان مفاهیمی چون «قصد» و «علم» را به نحو حقوقی به سامانه های هوشمند، ولو به طور مجازی یا استعاری، نسبت داد یا باید همواره در جست و جوی انسان پشت سامانه بود (زند و همکاران، ۲۰۲۴)؟ سوم آن که، از حیث سیاست کیفری و جرم شناختی، در جایی که امکان شناسایی فاعل انسانی مشخص دشوار یا ناممکن است، آیا توسل به مدل های «مسئولیت مبتنی بر ریسک» یا «مسئولیت جمعی و سازمانی» و حتی طراحی رژیم های ویژه مسئولیت کیفری برای اشخاص حقوقی و متصدیان فناوری، می تواند به نحو مشروع و کارآمد، خلأ موجود را پر کند؟

این پرسش ها، ضرورت بازنگری در مبانی نظری و سازوکارهای عملی حقوق کیفری را بیش از پیش آشکار می سازد. در این میان، نباید از ظرفیت ها و فرصت هایی که خود هوش مصنوعی برای نظام عدالت کیفری فراهم می کند غافل شد. در کنار مخاطرات و تهدیدها، هوش مصنوعی می تواند در کشف جرم، تحلیل داده های عظیم، پیش بینی نقاط جرم خیز، شناسایی شبکه های مجرمانه، پایش تراکنش های مالی مشکوک و حتی در فرایند دادرسی، از طریق سامان دهی اسناد و کمک به قاضی در تحلیل پرونده، نقشی کارآمد ایفا کند (خلیلی پاچی و همکاران، ۲۰۲۵؛ کاهدی و همکاران، ۲۰۲۵). با این حال، به کارگیری این ابزارها، خود با نگرانی هایی درباره تبعیض الگوریتمی، خطای سیستمی، نقض حریم خصوصی و تضعیف تضمین های دادرسی عادلانه همراه است (خلیلی پاچی و همکاران، ۲۰۲۵؛ کریمی، ۲۰۲۵). بدین ترتیب، همان گونه که هوش مصنوعی می تواند در خدمت ارتکاب جرم و گریز از مسئولیت قرار گیرد، می تواند، اگر به نحو صحیح تنظیم و نظارت شود، در خدمت پیشگیری و کشف جرم و تحقق عدالت کیفری نیز قرار گیرد؛ اما در هر دو صورت، فقدان چارچوب های روشن حقوقی، خطر سوءاستفاده، تبعیض و بی عدالتی را تشدید می کند (سردارنیا و همکاران، ۲۰۱۹).

با لحاظ مطالب فوق، هدف مقاله حاضر آن است که با رویکردی مروری - تحلیلی، مهم ترین چالش های حقوق کیفری در مواجهه با جرائم مبتنی بر هوش مصنوعی را شناسایی، دسته بندی و تحلیل نموده و در پرتو تجربیات نظام های حقوقی دیگر و مبانی فقهی و حقوقی ایران، تصویری منسجم تر از وضعیت موجود و نیازهای تقنینی و سیاست گذاری ارائه کند (میرلوحی و همکاران، ۲۰۲۵؛ الذبحاوی و همکاران، ۲۰۲۵). پرسش های محوری این پژوهش عبارتند از: (۱) جرائم مبتنی بر هوش مصنوعی چه ویژگی هایی دارند که آن ها را از جرائم سایبری سنتی متمایز می کند و این ویژگی ها چه پیامدهایی برای عناصر تشکیل دهنده جرم و انتساب مسئولیت کیفری دارد؟ (۲) مهم ترین خلأها و چالش های تقنینی، رویه ای و اثباتی حقوق کیفری ایران در مواجهه با این جرائم کدام است؟ (۳) چه الگوها و تجربیات تطبیقی ای در سطح بین المللی و منطقه ای قابل بهره برداری است و چگونه می توان آن ها را با مبانی فقه امامیه و ساختار حقوقی ایران سازگار ساخت (صمدیان و همکاران، ۲۰۲۵؛ خوشان تاش و همکاران، ۲۰۲۵)؟ روش پژوهش، کتابخانه ای و توصیفی - تحلیلی است و بر مطالعه نظام مند آثار و مقالات علمی موجود در حوزه حقوق کیفری، فقه، فلسفه حقوق و مطالعات تطبیقی، از جمله پژوهش های داخلی اخیر در زمینه مسئولیت کیفری هوش مصنوعی، خودروهای خودران، پول شویی مبتنی بر رمزارز و صیانت از حریم خصوصی در فضاهای مجازی نوین، تکیه دارد (میرلوحی و همکاران، ۲۰۲۵؛ کاهدی و همکاران، ۲۰۲۵؛ کریمی، ۲۰۲۵). بر این اساس، در ادامه، ابتدا به تبیین مبانی نظری و مفهومی جرائم مبتنی بر هوش مصنوعی پرداخته، سپس چالش های اساسی حقوق کیفری ایران را در این زمینه بررسی و در نتیجه، با

بهره‌گیری از تجربه سایر نظام‌های حقوقی، راهکارهایی برای ارتقای کارآمدی سیاست جنایی در مواجهه با این دسته از جرائم پیشنهاد خواهد شد (خلیلی پاجی و همکاران، ۲۰۲۵؛ الذبحاوی و همکاران، ۲۰۲۵).



شکل ۱. مدل مفهومی تعامل میان جرائم مبتنی بر هوش مصنوعی و چالش‌های حقوق کیفری

شکل ۱ رابطه‌ی میان مؤلفه‌های اصلی چرخه ارتکاب و مواجهه حقوق کیفری با جرائم مبتنی بر هوش مصنوعی را نمایش می‌دهد. در این مدل، ابتدا انواع رفتارهای مجرمانه ناشی از یا مرتبط با سامانه‌های هوشمند قرار گرفته و سپس چالش‌های بنیادین حقوق کیفری - شامل ابهام در انتساب مسئولیت، دشواری در شناسایی عنصر روانی، و موانع اثباتی - به‌عنوان لایه‌ی میانی نشان داده شده است. در بخش نهایی، واکنش‌های حقوق کیفری (تقنینی، قضایی و اجرایی) و رویکردهای تطبیقی به‌عنوان مسیرهای پاسخ‌دهی نمایان شده‌اند. فلش‌ها بیانگر جریان تأثیرگذاری متقابل میان این سطوح هستند؛ به این معنا که پیشرفت فناوری نه تنها الگوهای ارتکاب جرم را تغییر می‌دهد، بلکه خود موجب پیچیدگی بیشتر در ارزیابی و طراحی واکنش‌های حقوق کیفری می‌شود. این مدل می‌کوشد ساختار کلی مباحث پژوهش را در قالبی یکپارچه و قابل فهم نشان دهد.

۲. مبانی نظری و مفهومی جرائم مبتنی بر هوش مصنوعی

ورود فناوری‌های هوش مصنوعی به عرصه‌های مختلف حیات بشری، موجب شکل‌گیری گونه‌ای نو از رفتارهای زیان‌بار شده که نه به‌طور کامل در چارچوب جرائم سایبری سنتی قابل طبقه‌بندی است و نه می‌توان آن‌ها را بدون لحاظ ویژگی‌های متمایز این

سامانه‌ها، با جرائم کلاسیک تحلیل کرد. از منظر مفهومی، هوش مصنوعی به مجموعه‌ای از سامانه‌ها، الگوریتم‌ها و شبکه‌های محاسباتی اطلاق می‌شود که توانایی تحلیل داده‌ها، یادگیری الگوها، تصمیم‌گیری مستقل و انجام خودکار وظایف پیچیده را دارند و در بسیاری از موارد، رفتار آن‌ها نتیجه تعامل پویا میان داده‌های ورودی، ساختار الگوریتمی و فرایندهای خودیادگیر است (زندى و همکاران، ۲۰۲۴). این خودمختاری نسبی، سبب می‌شود که خروجی سامانه‌ها در برخی موارد، از کنترل مستقیم یا حتی پیش‌بینی طراح و بهره‌بردار خارج شود و از این طریق، ساحت جدیدی از «خطر فناوری محور» در فضای اجتماعی ایجاد گردد؛ خطری که با مفاهیم سنتی تقصیر، قصد و اراده در حقوق کیفری، سازگار و هم‌پوشان نیست (میرلوحی و همکاران، ۲۰۲۵). از این منظر، ریشه بسیاری از پیچیدگی‌ها در مواجهه با جرائم مبتنی بر هوش مصنوعی، نه در ماهیت مجرمانه رفتار، بلکه در ماهیت فاعل و اصل «انتساب» است؛ مسئله‌ای که مباحثی چون مسئولیت طراح، تولیدکننده، بهره‌بردار و حتی مسئولیت مبتنی بر ریسک را در ادبیات معاصر حقوق کیفری برجسته کرده است (صمدیان و همکاران، ۲۰۲۵).

جرائم مبتنی بر هوش مصنوعی را می‌توان به‌طور کلی در دو دسته اصلی تفکیک کرد: نخست، جرائمی که در آن انسان از هوش مصنوعی به‌عنوان ابزار یا وسیله برای ارتکاب جرم استفاده می‌کند؛ مانند استفاده از الگوریتم‌های یادگیری ماشین برای نفوذ در سامانه‌های مالی، طراحی حملات سایبری پیچیده یا تولید محتوای عمیق جعلی با هدف کلاهبرداری، اخاذی یا بی‌اعتبارسازی اشخاص (کاهدی و همکاران، ۲۰۲۵؛ کریمی، ۲۰۲۵). در این موارد، نقش انسانی همچنان محوری است و هوش مصنوعی نقشی ابزاری دارد، هرچند پیچیدگی فنی آن، کشف و اثبات جرم را دشوار می‌کند. دوم، دسته‌ای از جرائم که در آن‌ها رفتار یا تصمیم خود سامانه هوش مصنوعی، مستقیماً منجر به وقوع نتیجه زیان‌بار می‌شود؛ مانند تصمیم اشتباه خودروهای خودران در شرایط بحرانی، نقص عملکرد ربات‌های صنعتی یا پزشکی، یا تصمیم‌گیری خودکار سامانه‌های تسلیحاتی بدون کنترل یا مداخله انسانی کافی (خوشان‌تاش و همکاران، ۲۰۲۵). در این دسته، مسئله به‌طور مستقیم به پرسش از «فاعل» بازمی‌گردد؛ یعنی اینکه آیا می‌توان نتیجه را به انسان معینی نسبت داد یا باید ساختارهای جدیدی از مسئولیت را شناسایی کرد (زندى و همکاران، ۲۰۲۴). این وضعیت موجب شده است که پژوهشگران، هوش مصنوعی را نه فقط به‌عنوان ابزار، بلکه به‌عنوان «کنشگر فناورانه» تحلیل کنند؛ هرچند اکثریت نظام‌های حقوقی همچنان بر نفی «شخصیت حقوقی» برای سامانه‌های هوش مصنوعی تأکید دارند (الذبحاوی و همکاران، ۲۰۲۵).

تمایز جرائم مبتنی بر هوش مصنوعی با جرائم سایبری سنتی، از چند جهت اهمیت دارد. نخست آن‌که در جرائم سایبری معمول، ماهیت رفتار غالباً به مهارت انسانی و قصد مستقیم فاعل وابسته است، اما در جرائم مبتنی بر هوش مصنوعی، نقش سامانه و غیرقابل پیش‌بینی بودن برخی خروجی‌ها، می‌تواند موجب پیچیده شدن رابطه علیت و دشوار شدن تعیین نقطه «کنترل انسانی مؤثر» گردد (میرلوحی و همکاران، ۲۰۲۵). دوم آن‌که، در جرائم مبتنی بر هوش مصنوعی، داده‌ها نقش دوگانه‌ای دارند: هم «موضوع» و هم «ابزار» جرم هستند. به عبارتی، داده‌ها و الگوریتم‌ها، نه تنها امکان ایجاد رفتار مجرمانه را فراهم می‌کنند، بلکه در فرایند رسیدگی نیز به‌عنوان منبع اصلی اثبات یا رد ادعا تلقی می‌شوند و هرگونه دست‌کاری یا نقص در آن‌ها، می‌تواند عدالت کیفری را دچار اختلال جدی کند (خلیلی پاچی و همکاران، ۲۰۲۵). سوم آن‌که این جرائم، غالباً «چندلایه» و «چندکنشگری» هستند؛ یعنی ممکن است طراح الگوریتم در یک کشور، شرکت بهره‌بردار در کشور دیگر، داده‌های آموزشی از منبعی ثالث و قربانی در موقعیتی دیگر قرار داشته باشد؛ امری که بُعد فراملی جرائم مبتنی بر هوش مصنوعی را تقویت کرده و صلاحیت کیفری و همکاری‌های بین‌المللی را به چالش می‌کشد (عبدالهی و همکاران، ۲۰۲۵). چنین وضعیتی، ضرورت بازنگری در قواعد تعارض قوانین، صلاحیت قضایی و سازوکارهای تبادل داده را دوچندان می‌کند.

ظرفیت‌های مجرمانه هوش مصنوعی نیز، به موجب ساختار الگوریتمی و ماهیت یادگیری آن، با جرائم کلاسیک تفاوت دارد. برای مثال، الگوریتم‌های یادگیری عمیق، قابلیت ایجاد «الگوهای رفتاری مستقل» در پاسخ به شرایط جدید را دارند و همین امر، سبب می‌شود که در برخی موارد، سامانه رفتاری را بروز دهد که طراح آن هرگز به‌طور دقیق پیش‌بینی نکرده است (میرلوحی و همکاران، ۲۰۲۵). این مسئله، اهمیت تدوین استانداردهای شفافیت الگوریتمی، قابلیت حسابرسی، مستندسازی تصمیمات و نیز تکالیف گزارش‌دهی را دوچندان کرده است (الذبحاوی و همکاران، ۲۰۲۵). افزون بر این، از منظر جرم‌شناختی، استفاده مجرمانه از هوش مصنوعی، با ویژگی‌هایی چون سرعت، مقیاس‌پذیری، ناشناسی و قابلیت خودکارسازی همراه است؛ به‌گونه‌ای که یک عامل انسانی با حداقل دخالت، می‌تواند عملیات مجرمانه‌ای را در سطح گسترده، دقیق و چندلایه سازمان‌دهی کند (عبدالهی و همکاران، ۲۰۲۵). برای نمونه، در حوزه پول‌شویی، الگوریتم‌های هوشمند می‌توانند تراکنش‌های رمزآزری را به‌گونه‌ای طبقه‌بندی و پراکنده کنند که شناسایی منشأ و مقصد وجوه، حتی با ابزارهای نظارتی پیشرفته نیز دشوار شود (کاهدی و همکاران، ۲۰۲۵). از سوی دیگر، در حوزه حریم خصوصی و جرم‌انگاری تعرضات دیجیتال، هوش مصنوعی این امکان را فراهم می‌کند که اطلاعات زیستی، هویتی و رفتاری افراد در مقیاسی بسیار گسترده‌تر از ابزارهای سنتی جمع‌آوری و تحلیل شود و این امر، دامنه نقض حریم خصوصی را افزایش و ابزارهای مقابله را تضعیف می‌کند (کریمی، ۲۰۲۵). در مجموع، مبانی نظری و مفهومی جرائم مبتنی بر هوش مصنوعی نشان می‌دهد که این جرائم، صرفاً نسخه پیشرفته‌ای از جرائم سایبری نیستند، بلکه گونه‌ای جدید از رفتارهای زیان‌بارند که در آن‌ها، نقش سامانه، داده، الگوریتم و طراحی فنی، به اندازه یا حتی بیش از نقش عامل انسانی اهمیت دارد. این ویژگی‌ها سبب می‌شود که حقوق کیفری، در مواجهه با این دسته از جرائم، با پرسش‌هایی بنیادین در حوزه عناصر جرم، رابطه علیت، انتساب مسئولیت و مبانی ضمان مواجه شود و ضرورت بازنگری در ساختار مقررات، استانداردها و سیاست‌های جنایی بیش‌ازپیش احساس گردد (صمدیان و همکاران، ۲۰۲۵؛ خوشان‌تاش و همکاران، ۲۰۲۵). روشن است که در ادامه، تحلیل این چالش‌ها نیازمند بررسی عمیق‌تر جنبه‌های نظری و تطبیقی است که در بخش‌های بعدی به آن پرداخته خواهد شد.

۳. چالش‌های اساسی حقوق کیفری در مواجهه با جرائم مبتنی بر هوش مصنوعی

تحول سریع فناوری‌های هوش مصنوعی، حقوق کیفری را با چالش‌هایی بنیادین روبه‌رو کرده است؛ چالش‌هایی که نه تنها به سطح اجرا و سیاست‌گذاری محدود نیست، بلکه ساختار مفهومی و هنجاری این شاخه از حقوق را نیز تحت تأثیر قرار می‌دهد.

۱. نخستین چالش، مسئله «انتساب مسئولیت کیفری» است. ماهیت خودیادگیر و نیمه‌مستقل سامانه‌های هوش مصنوعی موجب می‌شود که در بسیاری موارد، رفتار زیان‌بار نتیجه ترکیبی از طراحی الگوریتمی، کیفیت داده‌های آموزشی، تصمیم‌های عملیاتی سامانه و نحوه بهره‌برداری انسان باشد. تداخل این عوامل، احراز رابطه علیت را دشوار کرده و تعیین فاعل مسئول را به یکی از پیچیده‌ترین مسائل حقوق کیفری معاصر تبدیل کرده است (میرلوحی و همکاران، ۲۰۲۵). هنگامی که خروجی سامانه از کنترل یا پیش‌بینی طراح خارج می‌شود، پرسش اساسی این است که آیا باید طراح را به‌سبب ضعف در طراحی، بهره‌بردار را به‌سبب عدم نظارت کافی، یا هیچ‌یک را به‌سبب عدم پیش‌بینی‌پذیری دقیق، مسئول شناخت. ادبیات جدید پیشنهادهایی چون «مسئولیت مبتنی بر ریسک»، «مسئولیت جمعی» و «مسئولیت لایه‌ای» را مطرح کرده است، اما هیچ‌یک هنوز به‌طور کامل جایگاهی قطعی در نظام‌های کیفری نیافته‌اند (صمدیان و همکاران، ۲۰۲۵).

۲. چالش دوم، مسئله «ابهام در عناصر جرم» است. عنصر مادی جرائم مبتنی بر هوش مصنوعی، برخلاف جرائم سنتی، اغلب در قالب یک کنش فیزیکی قابل رؤیت نیست و ممکن است صرفاً در سطح داده، الگوریتم یا تصمیم نرم افزاری رخ دهد. این امر موجب می شود که تحقق یا عدم تحقق رفتار مجرمانه، وابسته به تحلیل های پیچیده فنی و تخصصی باشد و بسیاری از نظام های کیفری فاقد معیارهای روشن برای تشخیص اینکه «رفتار» در معنای حقوق کیفری دقیقاً در چه نقطه ای از فرایند الگوریتمی رخ داده است، باشند (کاهدی و همکاران، ۲۰۲۵). از سوی دیگر، عنصر روانی جرم نیز با دشواری های اساسی مواجه است؛ زیرا در برخی موارد، قصد و اراده انسانی در لحظه وقوع رفتار زیان بار حضور ندارد و تصمیم مجرمانه به طور خودکار توسط سامانه اتخاذ می شود (زندى و همکاران، ۲۰۲۴). این وضعیت، بنیان سنتی «قصد» و «علم» را با پرسش مواجه کرده و نیاز به مدل های جدیدی از احراز مسئولیت ذهنی را برجسته ساخته است؛ مدلهایی که بتوانند بین خطای انسانی در طراحی، خطای داده ای در آموزش و رفتار مستقل سامانه تفکیک قائل شوند (خوشان تاش و همکاران، ۲۰۲۵).

۳. چالش سوم، «کشف جرم، جمع آوری ادله و فرآیند اثبات» است. ساختار پیچیده و چندلایه الگوریتم ها، استفاده گسترده از داده های رمزنگاری شده، و عدم شفافیت کافی در برخی مدل ها سبب می شود که فرآیند تشخیص منشأ رفتار زیان بار، بسیار دشوار باشد. در جرائم مبتنی بر هوش مصنوعی، بخش عمده ادله در قالب داده های فنی، گزارش های سیستمی، لاگ های نرم افزاری و ساختارهای الگوریتمی است و این امر نیازمند مهارت تخصصی در تحلیل داده و در مواردی، نیازمند سطحی از دسترسی است که ممکن است شرکت های فناوری به دلایل تجاری یا امنیتی از ارائه آن امتناع کنند (خلیلی پاچی و همکاران، ۲۰۲۵). علاوه بر این، تغییرپذیری سریع داده ها در سامانه های هوشمند، امکان از بین رفتن یا بازنویسی خودکار برخی داده ها را افزایش می دهد و همین مسئله چالش های جدی در حفظ زنجیره ادله ایجاد می کند. حتی در صورت دسترسی، عدم شفافیت الگوریتمی (black-box) می تواند مانع فهم قضایی از رابطه میان داده و نتیجه شود و به این ترتیب، عدالت کیفری را دچار اختلال سازد (عبدالهی و همکاران، ۲۰۲۵).

۴. چالش چهارم، «مرزهای صلاحیت کیفری و ابعاد فراملی» است. بسیاری از سامانه های هوش مصنوعی مبتنی بر داده های جهانی، پردازش ابری و مقررات متفاوت ملی عمل می کنند. در نتیجه، ممکن است طراحی الگوریتم در یک کشور، اجرای سامانه در کشوری دیگر و وقوع نتیجه زیان بار در جغرافیایی متفاوت باشد. چنین پراکندگی جغرافیایی، تعیین صلاحیت قضایی، انتخاب قانون حاکم، همکاری بین المللی و تبادل داده را دچار پیچیدگی می کند (عبدالهی و همکاران، ۲۰۲۵). علاوه بر این، برخی کشورها رویکردهای متفاوتی در خصوص تنظیم گری هوش مصنوعی اتخاذ کرده اند و این اختلاف، خلأهایی ایجاد می کند که مجرمان می توانند از آن ها برای فرار از پاسخگویی استفاده کنند. برای نمونه، عدم هماهنگی در مقررات مربوط به داده ها و حریم خصوصی، امکان سوءاستفاده از داده های زیستی و رفتاری افراد در کشورهایی با نظارت ضعیف را افزایش می دهد و آثار آن در کشورهایی با استانداردهای سخت گیرانه تر ظاهر می شود (کریمی، ۲۰۲۵).

۵. چالش پنجم، «نظام مندی و پیش بینی پذیری سیاست جنایی» است. سرعت توسعه هوش مصنوعی بسیار بیشتر از سرعت قانون گذاری است و همین امر سبب شده است که بسیاری از رفتارهای مجرمانه نوظهور، فاقد چارچوب حقوقی دقیق باشند. سیاست گذاران در بسیاری از کشورها، هنوز در مرحله تدوین استانداردهای پایه ای مانند مسئولیت تولیدکنندگان، الزامات شفافیت الگوریتمی و نظام های ارزیابی ریسک هستند، در حالی که سطوح پیچیده تری از جرائم الگوریتم محور در حال بروز است (الذبحاوی و همکاران، ۲۰۲۵). این شکاف میان سرعت فناوری و کندی قانون گذاری، خطر شکل گیری «مناطق خاکستری» را ایجاد می کند؛ مناطقی که در آن ها رفتارهای بالقوه زیان بار فاقد ضمانت اجرای کافی هستند و همین مسئله می تواند اعتماد عمومی به عدالت کیفری را کاهش دهد (میرلوحی و همکاران، ۲۰۲۵).

در مجموع، چالش های یادشده نشان می دهد که جرائم مبتنی بر هوش مصنوعی صرفاً مسئله ای فنی یا عملیاتی نیست، بلکه موضوعی ساختاری و بنیادین در حقوق کیفری معاصر است. این چالش ها نه تنها نیازمند اصلاح قوانین، بلکه مستلزم تحول در پارادایم های مفهومی مانند قصد، تقصیر، علیت، انتساب و ادله دیجیتال هستند. بنابراین، مواجهه مؤثر با جرائم مبتنی بر هوش مصنوعی، نیازمند ترکیبی از اصلاحات قانون گذاری، ارتقای ظرفیت های فنی، توسعه همکاری های بین المللی و بازنگری در اصول سنتی حقوق کیفری است؛ امری که در بخش های بعدی، از جمله رویکردهای تطبیقی و یافته های پژوهش، بیشتر مورد واکاوی قرار خواهد گرفت.

۴. رویکردهای تطبیقی نظام های حقوقی در مواجهه با جرائم مبتنی بر هوش مصنوعی

گسترش کاربردهای هوش مصنوعی در حوزه های اقتصادی، اجتماعی و امنیتی موجب شده است که بسیاری از نظام های حقوقی جهان در سال های اخیر به تدوین سیاست ها و چارچوب های حقوقی جدید برای مدیریت مخاطرات آن روی آورند. بررسی تطبیقی این رویکردها نشان می دهد که هرچند هنوز اجماع کاملی درباره شیوه مواجهه کیفری با جرائم مبتنی بر هوش مصنوعی شکل نگرفته است، اما گرایش های مشترکی در میان نظام های حقوقی مختلف قابل مشاهده است. یکی از مهم ترین این گرایش ها، حرکت از رویکرد صرفاً واکنشی به سمت رویکرد پیشگیرانه و تنظیم گرانه است. در بسیاری از کشورها، قانون گذاران تلاش کرده اند با وضع مقرراتی درباره شفافیت الگوریتمی، ارزیابی ریسک سامانه های هوش مصنوعی و تعیین مسئولیت برای تولیدکنندگان و بهره برداران، از وقوع آسیب های بالقوه جلوگیری کنند (الذبحاوی و همکاران، ۲۰۲۵). در این چارچوب، مسئولیت کیفری به عنوان آخرین ابزار مداخله مورد توجه قرار می گیرد و در کنار آن، ابزارهای نظارتی، استانداردهای فنی و سازوکارهای نظارت اداری نیز توسعه یافته اند.

در اتحادیه اروپا، رویکرد غالب مبتنی بر تنظیم گری مبتنی بر ریسک است. مقررات پیشنهادی و سپس تصویب شده در حوزه هوش مصنوعی در این منطقه، سامانه های هوش مصنوعی را بر اساس سطح خطر آن ها طبقه بندی می کند و برای هر سطح از ریسک، الزامات متفاوتی در نظر می گیرد. به عنوان مثال، سامانه هایی که در حوزه هایی مانند زیرساخت های حیاتی، عدالت کیفری، سلامت یا حمل و نقل استفاده می شوند، در دسته سامانه های پرخطر قرار می گیرند و باید الزامات سخت گیرانه ای مانند شفافیت، قابلیت حسابرسی و نظارت انسانی را رعایت کنند (عبدالهی و همکاران، ۲۰۲۵). در صورت عدم رعایت این الزامات، امکان اعمال ضمانت اجراهای کیفری یا شبه کیفری علیه شرکت ها و توسعه دهندگان وجود دارد. این رویکرد نشان می دهد که قانون گذاران اروپایی تلاش دارند با تمرکز بر مسئولیت تولیدکنندگان و ارائه دهندگان فناوری، از بروز آسیب های ناشی از عملکرد سامانه های هوش مصنوعی پیشگیری کنند. علاوه بر این، در نظام حقوقی اروپا، تأکید ویژه ای بر حفاظت از داده های شخصی و حریم خصوصی وجود دارد و مقررات مرتبط با داده ها به طور مستقیم بر نحوه طراحی و استفاده از سامانه های هوش مصنوعی تأثیر می گذارد (کریمی، ۲۰۲۵).

در ایالات متحده آمریکا، رویکردی متفاوت و نسبتاً انعطاف پذیرتر مشاهده می شود. در این کشور، به جای وضع یک قانون جامع و یکپارچه برای هوش مصنوعی، تنظیم گری عمدتاً از طریق ترکیبی از قوانین موجود، دستورالعمل های نهادی و سیاست های بخشی انجام می شود. بسیاری از جرائم مرتبط با سوء استفاده از هوش مصنوعی در چارچوب قوانین موجود مانند جرائم سایبری، کلاهبرداری رایانه ای، نقض حریم خصوصی یا جرائم مالی مورد رسیدگی قرار می گیرند (میرلوحی و همکاران، ۲۰۲۵). این

رویکرد نشان دهنده تمایل نظام حقوقی آمریکا به حفظ انعطاف پذیری حقوقی و اجتناب از محدودیت های سخت گیرانه ای است که ممکن است نوآوری فناوری را کاهش دهد. با این حال، در سال های اخیر بحث هایی درباره ضرورت وضع قوانین خاص برای هوش مصنوعی و تعیین مسئولیت شرکت های فناوری نیز در این کشور افزایش یافته است، به ویژه در حوزه هایی مانند خودروهای خودران، سامانه های تصمیم گیری خودکار و تولید محتوای عمیق جعلی (deepfake) که می توانند آثار کیفی و اجتماعی گسترده ای داشته باشند (کاهدی و همکاران، ۲۰۲۵).

در برخی کشورهای آسیایی نیز رویکردهای متفاوتی در حال شکل گیری است. برای مثال، چین تلاش کرده است با ترکیب نظارت دولتی قوی و مقررات دقیق در حوزه داده ها و الگوریتم ها، کنترل بیشتری بر توسعه و کاربرد هوش مصنوعی اعمال کند. در این چارچوب، شرکت های فناوری موظف هستند اطلاعات مربوط به الگوریتم ها، داده های آموزشی و نحوه عملکرد سامانه های خود را در اختیار نهادهای نظارتی قرار دهند و در صورت بروز تخلفات یا آسیب های ناشی از عملکرد سامانه ها، مسئولیت های حقوقی و کیفی مشخصی متوجه آن ها خواهد بود (خوشان تاش و همکاران، ۲۰۲۵). این رویکرد نشان می دهد که برخی نظام های حقوقی تلاش دارند با اعمال نظارت پیشینی قوی، احتمال وقوع جرائم مبتنی بر هوش مصنوعی را کاهش دهند. در کنار این تجربه های بین المللی، بسیاری از نظام های حقوقی در حال بررسی این پرسش هستند که آیا باید برای سامانه های هوش مصنوعی نوعی شخصیت حقوقی مستقل در نظر گرفت یا خیر. هرچند برخی نظریه پردازان پیشنهاد کرده اند که اعطای شخصیت حقوقی محدود به سامانه های پیشرفته می تواند مسئله مسئولیت را تا حدی حل کند، اما اغلب نظام های حقوقی هنوز این دیدگاه را نپذیرفته اند و ترجیح می دهند مسئولیت را در چارچوب انسان ها و نهادهای حقوقی موجود، مانند تولیدکنندگان، توسعه دهندگان یا بهره برداران سامانه ها، تحلیل کنند (صمدیان و همکاران، ۲۰۲۵). این رویکرد با اصول بنیادین حقوق کیفری که مسئولیت را به عامل انسانی پیوند می دهد، سازگارتر تلقی می شود.

در مجموع، مطالعه تطبیقی نشان می دهد که اگرچه کشورها راهبردهای متفاوتی در مواجهه با جرائم مبتنی بر هوش مصنوعی اتخاذ کرده اند، اما چند محور مشترک در میان آن ها قابل شناسایی است: تأکید بر تنظیم گری پیشگیرانه، توسعه استانداردهای شفافیت و پاسخگویی الگوریتمی، تقویت مسئولیت تولیدکنندگان و بهره برداران فناوری، و تلاش برای هماهنگی مقررات در سطح بین المللی. این روندها نشان می دهد که حقوق کیفری در حال گذار از چارچوب های سنتی به سمت مدل های نوینی است که بتوانند پیچیدگی های فناوری های هوشمند را بهتر مدیریت کنند (الذبحاوی و همکاران، ۲۰۲۵؛ عبدالمهدی و همکاران، ۲۰۲۵). بررسی این رویکردهای تطبیقی، زمینه مناسبی برای استخراج یافته های پژوهش فراهم می کند که در بخش بعدی به آن پرداخته خواهد شد.

۵. یافته های پژوهش

بررسی نظام مند منابع پیش گفته و تحلیل تطبیقی دیدگاه های مطرح شده نشان می دهد که چالش های حقوق کیفری در مواجهه با جرائم مبتنی بر هوش مصنوعی را می توان در چند محور کلیدی طبقه بندی کرد: چالش های مفهومی و نظری، چالش های تقنینی و ساختاری، چالش های اثباتی و آیین دادرسی، و چالش های فراملی و سیاست جنایی. یافته های پژوهش بیانگر آن است که اغلب نویسندگان، بیش از آنکه بر خلأ مطلق قانون تأکید داشته باشند، بر «ناکفایتی چارچوب های سنتی» در تبیین و پاسخ گویی به رفتارهای الگوریتم محور تمرکز کرده اند (میرلوحی و همکاران، ۲۰۲۵؛ زندی و همکاران، ۲۰۲۴). به عبارت دیگر،

مسئله اصلی، نبود مقررہ نیست، بلکه عدم انطباق مفاهیم کلاسیک همچون قصد، تقصیر، مباشرت، معاونت و رابطه علیت با واقعیت های فناوریانه جدید است. همچنین یافته ها نشان می دهد که در اکثر پژوهش ها، راهکارهای پیشنهادی در سه سطح ارائه شده اند: بازتفسیر مفاهیم موجود، اصلاح و تکمیل قوانین، و حرکت به سمت مدل های ترکیبی مسئولیت مبتنی بر ریسک و نظارت پیشینی (صمدیان و همکاران، ۲۰۲۵؛ الذبحاوی و همکاران، ۲۰۲۵).

جدول ۱: طبقه بندی چالش های حقوق کیفری در جرائم مبتنی بر هوش مصنوعی

ردیف	نوع چالش	شرح چالش	پیامد حقوقی
۱	مفهومی-نظری	ابهام در فاعل جرم و انتساب رفتار سامانه	دشواری در اعمال مسئولیت کیفری سنتی
۲	عنصر روانی	فقدان قصد مستقیم انسانی در لحظه وقوع نتیجه	تزلزل مبنای تقصیر و سوءنیت
۳	رابطه علیت	پیچیدگی زنجیره تصمیم گیری الگوریتمی	دشواری اثبات سببیت
۴	اثبات و ادله	عدم شفافیت الگوریتمها (Black Box)	اختلال در کشف حقیقت قضایی
۵	فراملی	پراکندگی جغرافیایی طراحی، اجرا و نتیجه	تعارض صلاحیت ها و قوانین
۶	سیاست جنایی	کندی قانون گذاری نسبت به سرعت فناوری	ایجاد خلأهای تنظیمی و مناطق خاکستری

تحلیل جدول شماره ۱ نشان می دهد که چالش های حقوق کیفری در این حوزه صرفاً به یک بعد خاص محدود نیست، بلکه ساختار چندلایه ای دارد. مهم ترین یافته آن است که مسئله «انتساب» به عنوان محور مرکزی اکثر مشکلات مطرح می شود؛ زیرا تقریباً تمامی چالش های دیگر، از جمله عنصر روانی و رابطه علیت، به نوعی به همین موضوع بازمی گردند (میرلوحی و همکاران، ۲۰۲۵). علاوه بر این، عدم شفافیت الگوریتمی که در ادبیات به عنوان مسئله جعبه سیاه شناخته می شود، نه تنها فرآیند اثبات را دشوار می کند، بلکه حتی امکان ارزیابی میزان تقصیر طراح یا بهره بردار را نیز محدود می سازد (خلیلی پاچی و همکاران، ۲۰۲۵). از سوی دیگر، یافته ها نشان می دهد که بعد فراملی این جرائم، بیش از جرائم سایبری سنتی است؛ زیرا در اینجا نه تنها داده ها، بلکه طراحی و تصمیم گیری نیز ماهیتی پراکنده دارند (عبدالهی و همکاران، ۲۰۲۵).

جدول ۲: رویکردهای پیشنهادی در قبال مسئولیت کیفری

رویکرد	مبنای نظری	مزایا	چالش ها
مسئولیت مبتنی بر تقصیر سنتی	احراز خطای انسانی در طراحی یا نظارت	سازگار با اصول کلاسیک	دشواری اثبات در سامانه های خودیادگیر

احتمال تضعیف اصل شخصی بودن مسئولیت	تقویت حمایت از بزه‌دیدگان	پذیرش خطر ذاتی فناوری	مسئولیت مبتنی بر ریسک
پیچیدگی در تعیین سهم هر عامل	واقع‌بینانه و متناسب با ساختار فناوری	تقسیم مسئولیت میان طراح، تولیدکننده و بهره‌بردار	مسئولیت جمعی/ لایه‌ای
افزایش بار اداری و هزینه‌های فناوری	پیشگیری مؤثر	کنترل و ارزیابی ریسک پیش از بهره‌برداری	تنظیم‌گری پیشینی

مطابق جدول شماره ۲، یافته‌های پژوهش نشان می‌دهد که هیچ رویکردی به‌تنهایی پاسخ‌گوی تمام ابعاد مسئله نیست. رویکرد مبتنی بر تقصیر سنتی اگرچه با مبانی کلاسیک حقوق کیفری هماهنگ است، اما در مواردی که رفتار سامانه غیرقابل پیش‌بینی بوده یا ناشی از یادگیری مستقل است، کارآمدی محدودی دارد (زندى و همکاران، ۲۰۲۴). در مقابل، رویکرد مبتنی بر ریسک، با پذیرش خطر ذاتی فناوری‌های پیشرفته، حمایت بیشتری از بزه‌دیدگان فراهم می‌کند، اما ممکن است با اصل شخصی بودن مسئولیت و ضرورت احراز سوءنیت تعارض یابد (صمدیان و همکاران، ۲۰۲۵). مدل مسئولیت لایه‌ای، که در برخی تحلیل‌های تطبیقی پیشنهاد شده، تلاش می‌کند سهم هر یک از بازیگران زنجیره فناوری را بر اساس میزان کنترل و نقش آن‌ها تعیین کند و به نظر می‌رسد از منظر سیاست جنایی، واقع‌بینانه‌تر باشد (الذبحاوی و همکاران، ۲۰۲۵). در آخر، بسیاری از پژوهش‌ها بر این نکته تأکید دارند که تمرکز صرف بر پاسخ کیفری کافی نیست و باید تنظیم‌گری پیشینی و استانداردگذاری فنی به‌عنوان مکمل مسئولیت کیفری تقویت شود (عبداللهی و همکاران، ۲۰۲۵).

جدول ۳: آثار جرائم مبتنی بر هوش مصنوعی بر ارکان کلاسیک جرم

رکن جرم	تحول ایجادشده	نتیجه تحلیلی
رکن قانونی	ناکفایتی برخی عناوین سنتی	ضرورت بازنگری یا تفسیر موسع
رکن مادی	تبدیل کنش فیزیکی به تصمیم الگوریتمی	نیاز به تعریف جدید از «رفتار»
رکن معنوی	تضعیف نقش قصد مستقیم	گرایش به مدل‌های تقصیر ساختاری
رابطه علیت	چندعاملی و شبکه‌ای شدن سببیت	پذیرش نظریه‌های سببیت انعطاف‌پذیر
مسئولیت	گسترش به بازیگران فناورانه	حرکت به سوی مسئولیت ترکیبی

جدول شماره ۳ نشان می‌دهد که جرائم مبتنی بر هوش مصنوعی نه‌تنها در سطح اجرا، بلکه در سطح ارکان کلاسیک جرم نیز تحول ایجاد کرده‌اند. رکن مادی، که در حقوق سنتی مبتنی بر یک کنش عینی و قابل انتساب به شخص معین بود، اکنون می‌تواند در قالب یک تصمیم نرم‌افزاری یا خروجی الگوریتمی متجلی شود؛ امری که نیازمند بازتعریف مفهوم «رفتار» در حقوق کیفری است (کاهدی و همکاران، ۲۰۲۵). همچنین، رکن معنوی که بر سوءنیت فردی استوار است، در مواردی که تصمیم توسط سامانه اتخاذ می‌شود، جای خود را به نوعی «تقصیر ساختاری» می‌دهد که به ضعف در طراحی، آموزش یا نظارت مربوط می‌شود (خوشان‌تاش و همکاران، ۲۰۲۵). یافته‌ها نشان می‌دهد که بسیاری از پژوهشگران بر ضرورت پذیرش نظریه‌های انعطاف‌پذیرتر در باب سببیت و مسئولیت تأکید دارند تا بتوانند ساختار شبکه‌ای و چندعاملی این جرائم را تحلیل کنند (میرلوحی

و همکاران، ۲۰۲۵). در جمع‌بندی یافته‌ها می‌توان گفت که مهم‌ترین نتیجه پژوهش آن است که حقوق کیفری در مواجهه با جرائم مبتنی بر هوش مصنوعی، نیازمند «تحول تدریجی و هوشمندانه» است، نه گسست کامل از مبانی سنتی. اکثریت منابع بر این باورند که می‌توان با بازتفسیر مفاهیم کلاسیک، تکمیل خلأهای تقنینی، توسعه استانداردهای شفافیت الگوریتمی و تقویت همکاری‌های بین‌المللی، چارچوبی متوازن میان حمایت از نوآوری و تضمین امنیت اجتماعی ایجاد کرد (الذبحاوی و همکاران، ۲۰۲۵؛ عبدالمهدی و همکاران، ۲۰۲۵). این یافته‌ها مبنای ارائه پیشنهادها نهایی در بخش نتیجه‌گیری خواهد بود.

۶. نتیجه‌گیری و پیشنهادات

گسترش سریع فناوری‌های مبتنی بر هوش مصنوعی، ساختارهای سنتی حقوق کیفری را با چالش‌های قابل توجهی مواجه کرده است. بررسی مباحث نظری، چالش‌های موجود و رویکردهای تطبیقی نشان می‌دهد که مهم‌ترین مسئله در این حوزه، ناتوانی کامل حقوق کیفری در مواجهه با این پدیده نیست، بلکه عدم انطباق برخی مفاهیم کلاسیک با واقعیت‌های جدید فناورانه است. در نظام سنتی حقوق کیفری، انتساب رفتار مجرمانه به یک شخص انسانی که دارای اراده و قصد است، محور اصلی تحلیل مسئولیت محسوب می‌شود؛ در حالی که در بسیاری از موارد مرتبط با هوش مصنوعی، نتیجه زیان‌بار ناشی از فرآیندهای پیچیده الگوریتمی، یادگیری ماشینی و تعامل داده‌هاست که به‌طور مستقیم قابل انتساب به یک اراده انسانی مشخص نیست. همین مسئله موجب شده است که مفاهیمی همچون فاعل جرم، سوءنیت، رابطه علیت و حتی تعریف رفتار مجرمانه با ابهامات جدیدی روبه‌رو شوند.

یافته‌های این پژوهش نشان می‌دهد که جرائم مبتنی بر هوش مصنوعی غالباً در بستری چندعاملی و شبکه‌ای رخ می‌دهند؛ به‌گونه‌ای که در شکل‌گیری آن‌ها مجموعه‌ای از بازیگران شامل طراحان الگوریتم، توسعه‌دهندگان نرم‌افزار، شرکت‌های ارائه‌دهنده خدمات، کاربران و حتی ساختارهای داده‌ای نقش دارند. در چنین شرایطی، تکیه صرف بر الگوهای سنتی مسئولیت فردی نمی‌تواند پاسخگوی تمام ابعاد این جرائم باشد. از سوی دیگر، تجربه نظام‌های حقوقی مختلف نشان می‌دهد که رویکردهای نوین بیشتر بر ترکیبی از مسئولیت حقوقی، تنظیم‌گری پیشینی و نظارت فناورانه تأکید دارند. به عبارت دیگر، سیاست جنایی مؤثر در این حوزه صرفاً مبتنی بر واکنش کیفری پس از وقوع جرم نیست، بلکه شامل سازوکارهای پیشگیرانه‌ای است که از طریق استانداردگذاری، ارزیابی ریسک و نظارت بر طراحی و بهره‌برداری از سامانه‌های هوش مصنوعی اعمال می‌شود. با توجه به نتایج به‌دست‌آمده، می‌توان گفت که آینده حقوق کیفری در مواجهه با جرائم مبتنی بر هوش مصنوعی در گرو نوعی تحول تدریجی در مفاهیم و سازوکارهای موجود است. این تحول باید به‌گونه‌ای صورت گیرد که از یک سو اصول بنیادین حقوق کیفری همچون شخصی بودن مسئولیت، قانونی بودن جرم و مجازات و لزوم احراز تقصیر حفظ شود و از سوی دیگر، امکان پاسخگویی مؤثر به مخاطرات فناوری‌های نوظهور فراهم گردد. در این راستا، بازتفسیر برخی مفاهیم سنتی و پذیرش مدل‌های انعطاف‌پذیرتر مسئولیت می‌تواند گامی مهم در جهت سازگار کردن حقوق کیفری با واقعیت‌های فناورانه باشد.

بر همین اساس، چند پیشنهاد قابل طرح است. نخست آنکه قانون‌گذاران باید با رصد مستمر تحولات فناوری، چارچوب‌های قانونی خود را به‌گونه‌ای به‌روزرسانی کنند که امکان پوشش رفتارهای مجرمانه نوظهور فراهم شود. دوم آنکه توسعه مقررات مربوط به شفافیت الگوریتمی و الزام شرکت‌های فناوری به ارائه اطلاعات درباره نحوه عملکرد سامانه‌های هوش مصنوعی می‌تواند نقش مهمی در تسهیل فرآیند کشف جرم و اثبات مسئولیت ایفا کند. سوم آنکه لازم است سازوکارهای مسئولیت چندلایه برای

بازیگران مختلف زنجیره فناوری در نظر گرفته شود تا سهم هر یک از طراحان، توسعه دهندگان و بهره برداران در صورت وقوع آسیب مشخص گردد. چهارم آنکه تقویت همکاری های بین المللی در حوزه مقابله با جرائم فناوریانه ضروری است؛ زیرا بسیاری از جرائم مبتنی بر هوش مصنوعی ماهیتی فرامرزی دارند و مقابله مؤثر با آنها مستلزم هماهنگی میان نظام های حقوقی مختلف است. سرانجام می توان گفت که هوش مصنوعی در عین حال که فرصت های گسترده ای برای پیشرفت علمی و اقتصادی فراهم می کند، مخاطرات جدیدی نیز برای نظام های حقوقی ایجاد کرده است. مواجهه مؤثر با این مخاطرات مستلزم رویکردی متوازن است که ضمن حمایت از نوآوری و توسعه فناوری، از امنیت اجتماعی و حقوق شهروندان نیز به طور مؤثر حمایت کند. در چنین چارچوبی، حقوق کیفری می تواند با تطبیق تدریجی خود با تحولات فناوریانه، همچنان نقش اساسی خود را در حفظ نظم و عدالت اجتماعی ایفا نماید.

منابع

- میرلوحی. (۲۰۲۵). چالش اثبات عنصر معنوی جرم در جرایم ارتكابی توسط هوش مصنوعی: امکان سنجی گذار از مسئولیت شخص محور به مسئولیت مبتنی بر ریسک مبتنی بر فقه و حقوق کیفری. پژوهش های حقوق جزا و جرم شناسی.
- زندی، رفیعی علوی، سید احسان. (۲۰۲۴). امکان سنجی مسئولیت کیفری در هوش مصنوعی بر پایه مبانی فلسفی آن. فلسفه حقوق، ۱۳(۱)، ۸۷-۱۰۶.
- خلیلی پاجی، عارف، احمدی. (۲۰۲۵). کاربرد هوش مصنوعی در حقوق کیفری بین المللی؛ از مسئولیت گریزی تا عدالت الگوریتمی. پژوهش های حقوقی.
- لیدا صمدیان، صادق مرادی، امیر سماواتی پیروز. (۲۰۲۵). مسئولیت کیفری در عصر هوش مصنوعی و جرائم سایبری؛ با تحلیل تطبیقی مبانی فقه امامیه و مالکی. پژوهش های تطبیقی فقه، حقوق و سیاست، ۱-۲۰.
- خوشان تاش، هیلگندورف، اریک. (۲۰۲۵). قانون گذاری خودروهای خودران: الگوبرداری از تجربه آلمان و چالش های پیش روی ایران. پژوهش های حقوق تطبیقی، ۲۹(۴)، ۳۴-۸۴.
- عبدالهی، پورات، صفایی. (۲۰۲۵). راهبردهای بازدارندگی سپاه در مواجهه با تهدیدات نوین امنیتی و مدیریت جرایم تروریستی در بستر نبردهای معاصر. مطالعات دفاع مقدس و نبردهای معاصر، ۷(۱۲)، ۲۳۹-۲۶۴.
- الذبحاوی، نبیل مهدی، بحر کاظمی، محمودی، باقری. (۲۰۲۵). تحلیل تطبیقی مقررات کیفری مرتبط با هوش مصنوعی و وسایل نقلیه خودران: مطالعه ای بر نظام های حقوقی سنگاپور، فرانسه و ایران. پژوهش تطبیقی حقوق اسلام و غرب، ۱۲(۱)، ۵۳-۷۸.
- کاهدی، شهربانو، غلامی، پوریا. (۲۰۲۵). کاربست هوش مصنوعی جهت مقابله با پولشویی از طریق رمزداری ها. مطالعات فقه و حقوق رسانه.
- رضوی فرد، موسوی، سید نعمت اله. (۲۰۱۶). مسئولیت کیفری در فضای سایبر در حقوق ایران. فصلنامه پژوهش حقوق کیفری، ۱۶(۱)، ۲۹-۴۵.
- سردارنیا، خلیل اله، سالاری شهر بابکی، میرزا مهدی. (۲۰۱۹). واکاوی پیشگیری از جرایم در ایران با تمرکز بر نظریه سیستمی در علوم سیاسی. فصلنامه پژوهش حقوق کیفری، ۷(۲۷)، ۴۳-۷۴.
- کریمی. (۲۰۲۵). صیانت از حریم خصوصی افراد در متاورس: راهکارهای مقابله با تعرض و سوءاستفاده. فصلنامه تحقیقات حقوقی، ۲۸(۴)، ۲۸۹-۳۰۵.