

پیامدهای اخلاقی پذیرش هوش مصنوعی (AI) در تصمیم‌گیری مالی

دکتر نوروز نوراله زاده

استادیار گروه حسابداری، واحد تهران جنوب، دانشگاه آزاد اسلامی، تهران، ایران.

(نویسنده مسئول)

nour547@yahoo.com

مرتضی صادقی

دانشجوی دکتری گروه حسابداری، واحد تهران جنوب، دانشگاه آزاد اسلامی، تهران، ایران.

Mortezasadeghi6271@gmail.com

چکیده

استفاده از هوش مصنوعی در بخش مالی، مسائل اخلاقی مهمی را ایجاد کرده که باید به آن‌ها رسیدگی شود. این مقاله به تحلیل پیامدهای اخلاقی استفاده از هوش مصنوعی در تصمیم‌گیری‌های مالی می‌پردازد و بر اهمیت یک چارچوب اخلاقی برای تضمین استقرار عادلانه و قابل اعتماد تأکید می‌کند. این مطالعه به بررسی ملاحظات اخلاقی مختلف، از جمله نیاز به بررسی و رفع سوگیری الگوریتمی، ترویج شفافیت و توضیح‌پذیری در سیستم‌های هوش مصنوعی، و رعایت مقرراتی که از عدالت، پاسخگویی و اعتماد عمومی محافظت می‌کنند، می‌پردازد. این مقاله با ترکیب تحقیقات و شواهد تجربی، رابطه پیچیده بین نوآوری هوش مصنوعی و یکپارچگی اخلاقی در امور مالی را برجسته می‌کند. برای مقابله با این موضوع، این مقاله یک چارچوب اخلاقی جامع و قابل اجرا ارائه می‌دهد که از دستورالعمل‌های روشن، ساختارهای حاکمیتی، ممیزی‌های منظم و همکاری بین ذینفعان حمایت می‌کند. هدف این چارچوب به حداکثر رساندن پتانسیل هوش مصنوعی ضمن به حداقل رساندن اثرات منفی و عواقب ناخواسته آن است. این مطالعه به عنوان یک منبع ارزشمند برای سیاست‌گذاران، متخصصان صنعت، محققان و سایر ذینفعان، تسهیل‌کننده بحث‌های آگاهانه، تصمیم‌گیری مبتنی بر شواهد و توسعه بهترین رویه‌ها برای ادغام مسئولانه هوش مصنوعی در بخش مالی می‌باشد. هدف نهایی تضمین عدالت، شفافیت و پاسخگویی ضمن بهره‌مندی از مزایای هوش مصنوعی برای هر دو بخش مالی و جامعه است.

واژه‌های کلیدی: هوش مصنوعی، اخلاق، شفافیت، پاسخگویی، امنیت.

مقدمه

هوش مصنوعی (AI) یک حوزه و فناوری نوظهور است که کاربردهایی زیادی در زمینه‌های مختلف، از جمله امور مالی، دارد. این فناوری در تصمیم‌گیری‌های آگاهانه در فضای مالی، با استفاده از حجم عظیمی از داده‌ها که با سرعتی باورنکردنی پردازش می‌شوند، به کار گرفته شده است. هوش مصنوعی بینش‌هایی را از همان داده‌ها کشف می‌کند که روش‌های سنتی مدل‌سازی ممکن است از دست بدهند، که باعث می‌شود مؤسسات مالی تصمیمات بهتر و سریع‌تری بگیرند. با این حال، پذیرش هوش مصنوعی در خدمات مالی و تصمیم‌گیری نیز ملاحظات اخلاقی مهمی را ایجاد می‌کند، زیرا مدل‌های یادگیری ماشین مسائل مربوط به سوگیری، شفافیت و حریم خصوصی را در تصمیم‌گیری ایجاد می‌کنند. در حالی که هوش مصنوعی مزایای رقابتی را برای شرکت‌ها و تصمیم‌گیرندگان اقتصادی به طور یکسان ارائه می‌دهد، نحوه مدیریت ریسک‌ها و رعایت استانداردهای اخلاقی برای پذیرش اخلاقی آن در تصمیمات مالی حیاتی است.

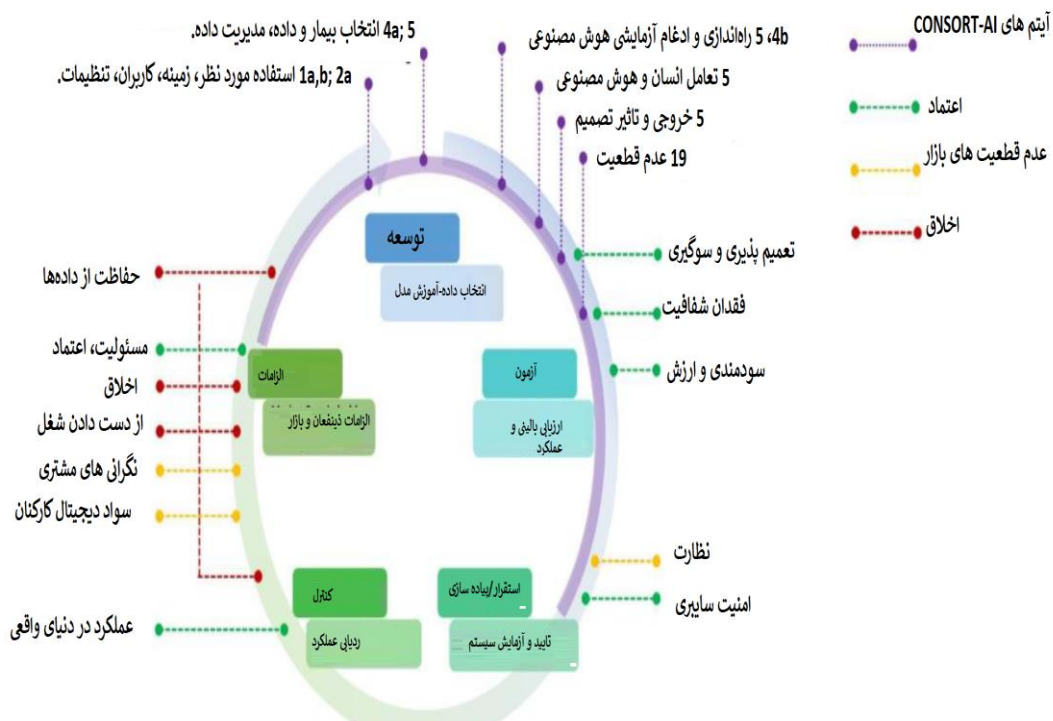
چارچوب‌های اخلاقی در تصمیم‌گیری‌های مالی

تصمیم‌گیری مالی یک فرآیند پیچیده است که به یک چارچوب اخلاقی قوی نیاز دارد تا افراد و سازمان‌ها را در اتخاذ تصمیماتی که نه تنها سودآور هستند، بلکه از نظر اخلاقی نیز سالم هستند، راهنمایی کند. چارچوب‌های اخلاقی در تصمیم‌گیری مالی مجموعه‌ای از اصول و دستورالعمل‌ها را ارائه می‌دهند که به افراد کمک می‌کند در شرایط پیچیده و مبهم مالی، مسیر درست را پیدا کنند. با رعایت استانداردهای اخلاقی، افراد می‌توانند اطمینان حاصل کنند که اقدامات آنها در راستای منافع همه ذینفعان درگیر، از جمله سهامداران، کارکنان و جامعه است. علاوه بر این، تصمیمات اخلاقی می‌تواند به ایجاد اعتماد و اطمینان در بازارهای مالی کمک کند و منجر به ثبات و پایداری بیشتر در بلندمدت شود (سایکانت^۱، ۲۰۲۴).

شفافیت و پاسخگویی در الگوریتم‌های هوش مصنوعی

شفافیت و پاسخگویی الگوریتم‌های هوش مصنوعی برای اطمینان از عادلانه و اخلاقی بودن تصمیمات مالی بسیار مهم هستند. با شفاف‌تر کردن الگوریتم‌های هوش مصنوعی، ذینفعان می‌توانند بهتر درک کنند که چگونه تصمیمات گرفته می‌شوند و هرگونه سوگیری که ممکن است ظاهر شود را شناسایی کنند (اولاتویه و همکاران^۲، ۲۰۲۴). یکی دیگر از اقدامات پاسخگویی، مسئول دانستن توسعه‌دهندگان و کاربران هوش مصنوعی در قبال نتایج تصمیمات آنها است، بنابراین تصمیم‌گیری اخلاقی و آگاهانه در بخش مالی ترویج می‌شود. بدون شفافیت و پاسخگویی، خطر تداوم سوگیری‌های موجود، تبعیض علیه گروه‌های خاص و ایجاد بی‌اعتمادی به سیستم‌های هوش مصنوعی که زیربنای تصمیم‌گیری‌های مالی هستند، وجود دارد.

شکل (۱): چرخه عمر هوش مصنوعی (AI) (شوندیک و کرویس^۳، ۲۰۲۱)



¹ Saikant

² Oletoye & et all

³ Schwendicke & et all

چرخه عمر برنامه‌های کاربردی هوش مصنوعی (AI) معمولاً با ارزیابی نیازها آغاز می‌شود، سپس توسعه، آزمایش، استقرار AI، نظارت و ارزیابی مجدد بالینی انجام می‌گیرد. جنبه‌های مختلفی در این چرخه عمر شناسایی شده‌اند که به‌عنوان موانعی برای پذیرش برنامه‌های کاربردی هوش مصنوعی عمل می‌کنند اعتماد به AI، شک و تردید بازار و نگرانی‌های اخلاقی (اماناس و همکاران، ۲۰۲۰): این موارد در چرخه عمر (با نقاط سبز، زرد و قرمز نشان داده شده‌اند) وجود دارند. موارد توسعه‌یافته‌ی استانداردهای گزارش‌دهی آزمایش‌های یکپارچه CONSORT-AI (مثلاً موارد ۱، a,b; 2a; 4a,b; 5; 19 که با خطوط قرمز مشخص شده‌اند) عمدتاً مربوط به مراحل توسعه و آزمایش هستند، به‌ویژه زمانی که آزمایش‌های کنترل‌شده‌ی تصادفی گزارش می‌شوند (نیم‌دایره‌ی بنفش).

سوگیری و عدالت در تصمیم‌گیری AI

در تصمیم‌گیری با استفاده از هوش مصنوعی، عدالت یک ملاحظه‌ی حیاتی است که باید برای به حداقل رساندن پیامدهای اخلاقی به آن پرداخته شود. یکی از نگرانی‌های اصلی این است که مجموعه‌های داده‌ی سوگیرانه می‌توانند در هنگام استفاده از الگوریتم‌های هوش مصنوعی در فرایندهای تصمیم‌گیری مالی، به نتایج تبعیض‌آمیز منجر شوند. تحقیقات نشان می‌دهند که داده‌های تاریخی مورد استفاده برای آموزش مدل‌های AI ممکن است انحرافات اجتماعی را منعکس کرده و منجر به رفتار ناعادلانه با گروه‌های جمعیتی خاص شوند. اگر تصمیمات وام‌دهی بر اساس شیوه‌های وام‌دهی گذشته باشد که به‌طور نامتناسب یک گروه را نسبت به گروه دیگر ترجیح می‌دهد، سیستم AI ممکن است این سوگیری را حفظ کند و منجر به نتایج ناعادلانه برای اقلیت شود. برای کاهش این ریسک، پیاده‌سازی مکانیسم‌هایی که شفافیت و پاسخگویی در سیستم‌های AI را تسهیل می‌کنند، ضروری است و این امر اجازه می‌دهد تا فرایندهای تصمیم‌گیری سوگیرانه شناسایی و حذف شوند. سازمان‌ها می‌توانند با اطمینان از اینکه الگوریتم‌ها به‌طور مداوم برای اعمال عدالت، پایش و حسابرسی می‌شوند، استانداردهای اخلاقی را برای ارتقای اعتماد به هوش مصنوعی مورد استفاده در بخش مالی حفظ کنند (کاستلنو^۱، ۲۰۲۴).

نگرانی‌های مربوط به حریم خصوصی و امنیت داده

نگرانی‌های مربوط به حریم خصوصی و امنیت داده هنگام بررسی پذیرش هوش مصنوعی در تصمیم‌گیری مالی بسیار مهم هستند. از آنجا که هوش مصنوعی حجم عظیمی از داده‌ها را جمع‌آوری و تجزیه و تحلیل می‌کند، این ریسک وجود دارد که اطلاعات حساس مورد سوء استفاده یا به خطر بیفتد. موسسات مالی باید اقدامات امنیتی قوی را برای محافظت از داده‌های مشتریان و حفظ اعتماد به سیستم ارائه دهند. نهادهای نظارتی باید دستورالعمل‌ها و استانداردهایی را برای اطمینان از استفاده اخلاقی از هوش مصنوعی در امور مالی و همزمان حفاظت از حقوق حریم خصوصی افراد وضع کنند (فابریگ و بوگونی^۲، ۲۰۲۳). اصول و دستورالعمل‌های هوش مصنوعی سازمان همکاری اقتصادی و توسعه (OECD) نیز ادعا می‌کنند که کیفیت و کفایت پردازش و نمایش داده‌ها با ریسک‌های خروجی مدل‌های گمراه‌کننده و مدل‌های نادرست یا غیرقابل اعتماد مرتبط است. چالش‌های حریم خصوصی داده‌ها در GenAI به دلیل تعداد زیاد مجموعه داده‌هایی که بر روی آنها آموزش دیده است و همچنین می‌تواند از هر منبع عمومی باشد، قابل توجه است. این داده‌ها احتمالاً شامل اطلاعات محافظت‌شده‌ی IP هستند، که ممکن است بدون مجوز مناسب باشند، و مشکلات اضافی در مورد اعتبار خروجی‌ها ایجاد می‌کنند. می‌توان از شیوه‌های متعالی برای مدیریت داده‌ها و شیوه‌های حکمرانی برای

¹ Castelnovo

² Fabrègue & Bogoni

اطمینان از کیفیت داده‌ها، کفایت داده‌ها بر اساس استفاده‌ی مورد نظر، حریم خصوصی داده‌ها هنگام گنجاندن داده‌های مشتریان مالی به مدل و اعتبار داده‌ها در زمانی که منبع/حقوق کپی مرتبط و مناسب اطلاع‌رسانی می‌شوند، استفاده کرد. در این حالت، رضایت آگاهانه می‌تواند برای این منظور به‌دست آید (پاواشه، کدم، زیرنگه و کتکار، ۲۰۲۳). محافظت از داده‌های فعال در حال انتقال و در حال استفاده برای حفظ محرمانگی، یکپارچگی و در دسترس بودن اطلاعات حیاتی بسیار مهم است (جیمز، ۲۰۲۳).

تأثیر بر اشتغال و نیروی کار

پذیرش هوش مصنوعی در فرایندهای تصمیم‌گیری مالی می‌تواند به‌طور قابل‌توجهی بر اشتغال و نیروی کار تأثیر بگذارد. هوش مصنوعی می‌تواند وظایفی را که به‌طور سنتی توسط انسان انجام می‌شده است، خودکار کند، در حالی که تکنیک‌های متداول مانند یادگیری ماشینی نیز کارایی را بهبود می‌بخشند. این کار ممکن است منجر به کاهش تقاضا برای نقش‌های شغلی خاص، به‌ویژه مواردی که شامل وظایف تکراری یا وظایف مبتنی بر قوانین هستند، شود. با این حال، مهم است که توجه داشته باشیم هوش مصنوعی مشاغل جدیدی را در زمینه‌هایی مانند تجزیه و تحلیل داده‌ها، یادگیری ماشینی و برنامه‌نویسی AI ایجاد خواهد کرد. با گسترش جهان در هوش مصنوعی و خدمات مالی، شرکت‌ها باید نیروی کار خود را ارزیابی مجدد کنند و برای تطبیق با تغییر چشم‌انداز صنعت خود آموزش ارائه دهند. تأثیر هوش مصنوعی بر مشاغل و نیروی کار به نحوه‌ی مدیریت گذار توسط سازمان‌ها و سرمایه‌گذاری در نوسازی نیروی کار برای آینده بستگی دارد (OECD، 2023).

چالش‌های نظارتی و مسائل انطباق

بکارگیری مدل‌های AI در امور مالی به‌دلیل ویژگی‌های منحصربه‌فرد فناوری AI، مانند عدم شفافیت، پیچیدگی و پتانسیل سوگیری، چالش‌هایی را ایجاد می‌کند. برای رسیدگی مؤثر به این چالش‌ها، استراتژی‌های همکاری با نهادهای نظارتی ضروری است. پیچیدگی‌های مقررات و انطباق مدل AI شامل عدم شفافیت و قابلیت توضیح مدل‌های AI است که می‌تواند به‌عنوان «جعبه‌های سیاه» عمل کند و فاقد شفافیت و پاسخگویی است. سوگیری و عدالت نیز نگرانی‌هایی هستند، زیرا مدل‌های AI می‌توانند سوگیری‌های موجود در داده‌های آموزشی را به ارث ببرند. علاوه بر این، ماهیت پویای هوش مصنوعی چالش‌هایی را برای چارچوب‌های نظارتی سنتی ایجاد می‌کند که ممکن است در تلاش برای همگامی با پیشرفت‌های فناوری باشند. استراتژی‌های همکاری با نهادهای نظارتی شامل تعامل و گفتگو برای افزایش درک متقابل، آزمایش‌های تحت نظارت، راهنمایی و استانداردهای استفاده‌ی مسئولانه از هوش مصنوعی و ایجاد ظرفیت از طریق برنامه‌های آموزشی و کارگاه‌ها است. همکاری و مشارکت بین مؤسسات مالی و نهادهای نظارتی برای پیمودن مؤثر پیچیدگی‌های مقررات و انطباق مدل AI در امور مالی بسیار مهم است. با همکاری یکدیگر، ذینفعان می‌توانند به چالش‌های نظارتی بپردازند، پذیرش مسئولانه‌ی هوش مصنوعی را ارتقا دهند و اطمینان حاصل کنند که با استانداردهای اخلاقی در تصمیم‌گیری مالی مطابقت دارد (Ejimofofor، اوولابی، ۲۰۲۳).

خطرات بالقوه‌ی پذیرش AI در تصمیم‌گیری مالی

¹ Jiménez

² Efijemue، Ejimofofor، Owolabi

پذیرش هوش مصنوعی در تصمیم‌گیری مالی مستلزم خطرات بالقوه‌ای است که برای اطمینان از استفاده‌ی اخلاقی و مسئولانه باید به دقت مدیریت شود. خطرات اصلی شفافیت، پاسخگویی و سوگیری هستند که بر اهمیت نظارت در توسعه و پیاده‌سازی AI برای بخش مالی تأکید می‌کنند. خطرات شفافیت شامل ماهیت غیرشفاف الگوریتم‌های AI است که درک فرایندهای تصمیم‌گیری را دشوار می‌کند و منجر به بی‌اعتمادی، مسائل نظارتی و چالش‌ها در توضیح تصمیمات می‌شود. ماهیت جعبه‌ی سیاه برخی از این مدل‌های هوش مصنوعی مانع از شناسایی خطاها، سوگیری‌ها یا رفتارهای غیر اخلاقی می‌شود. خطرات پاسخگویی از عدم اطمینان در مورد مسئولیت تصمیمات هدایت‌شده توسط AI ناشی می‌شود که ابهام در انتساب مسئولیت خطاها، سوگیری‌ها یا رفتارهای غیر اخلاقی و انطباق با الزامات قانونی و نظارتی را پیچیده می‌کند. خطرات سوگیری از الگوریتم‌های AI ناشی می‌شود که سوگیری‌ها را از داده‌های آموزشی به ارث می‌برند و ممکن است منجر به نتایج مغرضانه، شیوه‌های تبعیض‌آمیز، آسیب به شهرت و پیامدهای غیرعمدی مانند حفظ نابرابری یا محروم کردن گروه‌های خاص شود (ماندیچ و همکاران^۱، ۲۰۲۳).

هوش مصنوعی و معاملات الگوریتمی

پذیرش هوش مصنوعی (AI) در معاملات الگوریتمی بازارهای مالی را متحول کرده و امکان اجرای معاملات با سرعت بسیار زیاد و خودکارسازی استراتژی‌های پیچیده را فراهم کرده است. با این حال، این ترکیب قدرتمند همچنین ریسک‌های اخلاقی مهمی را ایجاد می‌کند که باید برای حفظ تمامیت و ثبات سیستم مالی به آنها پرداخته شود.

✓ تقویت نوسانات بازار و ریسک‌های سیستمیک: یکی از نگرانی‌های اصلی اخلاقی در مورد معاملات الگوریتمی هوش مصنوعی، پتانسیل آن در تقویت نوسانات بازار و تشدید ریسک‌های سیستمیک است. این الگوریتم‌های پیچیده می‌توانند با سرعت بسیار زیاد خرید و فروش کنند و باعث ایجاد سفارش‌های آبشاری شوند که می‌تواند بازارها را بی‌ثبات کند. ماهیت پر سرعت معاملات الگوریتمی و فقدان نظارت انسانی می‌تواند حلقه‌های بازخورد ایجاد کند، جایی که الگوریتم‌ها به معاملات یکدیگر واکنش نشان می‌دهند و ممکن است باعث نوسانات شدید قیمت و اختلالات بازار شوند.

✓ علاوه بر این، ارتباط متقابل بازارهای مالی و اتکا به مدل‌های هوش مصنوعی مشابه در میان موسسات مختلف می‌تواند منجر به ریسک‌های سیستمی شود. اگر الگوریتم‌های متعددی رفتار معاملاتی مشابهی از خود نشان دهند یا به طور یکسان به رویدادهای بازار واکنش نشان دهند، می‌تواند باعث اقدامات همزمان شود، نوسانات بازار را تشدید کند و به طور بالقوه منجر به بی‌ثباتی گسترده شود. "سقوط ناگهانی" سال ۲۰۱۰، که در آن میانگین صنعتی داو جونز در عرض چند دقیقه نزدیک به ۱۰۰۰ واحد سقوط کرد، اغلب به عنوان نمونه‌ای از پتانسیل معاملات الگوریتمی مبتنی بر هوش مصنوعی در دامن زدن به نوسانات شدید بازار ذکر می‌شود (کیریلنکو و همکاران^۲، ۲۰۱۷).

✓ فقدان نظارت و کنترل انسانی یکی دیگر از نگرانی‌های اخلاقی ناشی از فقدان نظارت و کنترل انسانی بر الگوریتم‌های معاملاتی هوش مصنوعی است. این سیستم‌های پیچیده به گونه‌ای طراحی شده‌اند که به طور خودکار عمل کنند و بر اساس حجم عظیمی از داده‌ها و مدل‌های پیچیده ریاضی تصمیمات آنی بگیرند (کیریلنکو و همکاران، ۲۰۱۳). در حالی که این استقلال، معاملات فوق‌العاده سریع را ممکن می‌سازد، سولاتی را نیز در مورد مسئولیت‌پذیری و شفافیت این الگوریتم‌ها مطرح می‌کند. ابهام مدل‌های هوش مصنوعی، که اغلب به عنوان "جعبه‌های سیاه" از آنها یاد می‌شود، می‌تواند درک کامل فرایندهای تصمیم‌گیری زیربنایی را برای انسان‌ها دشوار

¹ Mandych & et all

² Kirilenko & et all

کند. این فقدان شفافیت می‌تواند توانایی شناسایی و تصحیح سوگیری‌ها، خطاها یا پیامدهای ناخواسته احتمالی در الگوریتم‌ها را مختل کند (کیریلنکو و لو، ۲۰۱۳). علاوه بر این، سرعت بالایی که این الگوریتم‌ها با آن کار می‌کنند، می‌تواند مداخله و بازپس‌گیری کنترل را برای معامله‌گران یا تنظیم‌کننده‌های انسانی در صورت نقص یا رفتار غیرمنتظره دشوار کند.

✓ چالش‌های مسئولیت‌پذیری در صورت اختلالات بازار یا زیان‌های مالی ناشی از الگوریتم‌های معاملاتی هوش مصنوعی، تعیین مسئولیت به یک چالش مهم تبدیل می‌شود. ماهیت پیچیده این الگوریتم‌ها، همراه با دخالت طرف‌های متعدد (توسعه‌دهندگان، موسسات مالی، معامله‌گران و غیره)، تعیین مسئولیت را دشوار می‌کند. این فقدان مسئولیت‌پذیری روشن می‌تواند اعتماد عمومی به سیستم مالی را تضعیف کرده و اصول اخلاقی شفافیت و مسئولیت را تضعیف کند. علاوه بر این، ماهیت جهانی معاملات الگوریتمی، مسائل مربوط به صلاحیت قضایی را مطرح می‌کند، زیرا سیستم‌های هوش مصنوعی ممکن است در کشورهای متعدد با چارچوب‌های نظارتی متفاوت توسعه و مستقر شوند. این پراکندگی نظارتی می‌تواند شکاف‌هایی در نظارت و مسئولیت‌پذیری ایجاد کند و به طور بالقوه اجازه دهد تا شیوه‌های غیراخلاقی از بین شکاف‌ها عبور کنند.

پرداختن به این نگرانی‌های اخلاقی مستلزم یک رویکرد چندوجهی است که شامل همکاری بین موسسات مالی، نهادهای نظارتی و توسعه‌دهندگان هوش مصنوعی است. راه حل‌های بالقوه عبارتند از:

❖ ایجاد چارچوب‌های حاکمیتی قوی و استانداردهای صنعت برای توسعه و استقرار الگوریتم‌های معاملاتی هوش مصنوعی، با تمرکز بر شفافیت، مسئولیت‌پذیری و مدیریت ریسک (بروندیج و همکاران^۱، ۲۰۱۸).

❖ اجرای مکانیزم‌های نظارت انسانی و "کلیدهای توقف" که امکان مداخله و کنترل انسانی بر الگوریتم‌های معاملاتی هوش مصنوعی را در صورت نقص یا رفتار غیرمنتظره فراهم می‌کند.

❖ تقویت همکاری و هماهنگی نظارتی در بین حوزه‌های قضایی برای اطمینان از نظارت و اقدامات مسئولانه سازگار برای سیستم‌های معاملاتی الگوریتمی مبتنی بر هوش مصنوعی.

❖ ترویج حسابرسی‌های الگوریتمی و تست استرس برای شناسایی سوگیری‌ها، خطاها یا ریسک‌های سیستمی احتمالی مرتبط با الگوریتم‌های معاملاتی هوش مصنوعی و اجرای اقداماتی برای کاهش این خطرات (بروندیج و همکاران، ۲۰۱۸).

❖ تقویت شیوه‌های توسعه هوش مصنوعی اخلاقی، مانند گنجاندن اصول اخلاقی در طراحی و آموزش مدل‌های هوش مصنوعی و به کارگیری تیم‌های متنوع برای کاهش سوگیری‌ها و ترویج نوآوری مسئولانه (فلوریدا و همکاران^۲، ۲۰۱۹).

با پرداختن فعالانه به این نگرانی‌های اخلاقی و اجرای پادمان‌های مناسب، صنعت مالی می‌تواند از قدرت معاملات الگوریتمی مبتنی بر هوش مصنوعی بهره‌برد و در عین حال یکپارچگی، ثبات و استانداردهای اخلاقی سیستم مالی جهانی را حفظ کند.

تصمیم‌گیری اخلاقی در پیاده‌سازی هوش مصنوعی

تصمیم‌گیری اخلاقی در تضمین اینکه مزایای فناوری هوش مصنوعی به حداکثر می‌رسد و در عین حال آسیب‌های احتمالی به جامعه به حداقل می‌رسد، مهم است. یکی از ملاحظات کلیدی در این فرآیند، تقسیم مسئولیت بین انسان‌ها و

¹ Brundage & et all

² Floridi & et all

سیستم‌های هوش مصنوعی است. نقش‌ها و تعهدات انسان‌ها در حفظ سیستم‌های هوش مصنوعی باید به وضوح تعریف شود تا از خطاهای احتمالی که می‌تواند بر فرآیندهای تصمیم‌گیری تأثیر منفی بگذارد، جلوگیری شود. علاوه بر این، شفافیت و مسئولیت‌پذیری باید در طراحی و استقرار سیستم‌های هوش مصنوعی گنجانده شود تا اطمینان حاصل شود که اصول اخلاقی در طول فرآیند تصمیم‌گیری رعایت می‌شوند (بالاسوبرامانیان و همکاران^۱، ۲۰۲۴). در اینجا برخی از جنبه‌های کلیدی که باید در تصمیم‌گیری اخلاقی در پیاده‌سازی هوش مصنوعی در نظر گرفته شوند، آورده شده است:

متعادل‌سازی مسئولیت‌های انسان و هوش مصنوعی

نظارت انسانی: حفظ نظارت انسانی برای تضمین مسئولیت‌پذیری، شفافیت و ملاحظات اخلاقی ضروری است. انسان‌ها مسئول تعیین اهداف، طراحی الگوریتم‌های هوش مصنوعی، تفسیر نتایج و اتخاذ تصمیمات نهایی بر اساس اصول اخلاقی هستند.

قابلیت‌ها و محدودیت‌های هوش مصنوعی: درک قابلیت‌ها و محدودیت‌های سیستم‌های هوش مصنوعی برای تعیین میزان خودمختاری آن‌ها در تصمیم‌گیری بسیار مهم است. انسان‌ها باید زمانی که سیستم‌های هوش مصنوعی به محدودیت‌های خود می‌رسند یا با معضلات اخلاقی مواجه می‌شوند که نیازمند قضاوت انسانی است، مداخله کنند.

ادغام ملاحظات اخلاقی

چارچوب‌های اخلاقی: ایجاد چارچوب‌ها و دستورالعمل‌های اخلاقی روشن برای استقرار هوش مصنوعی برای اطمینان از اینکه سیستم‌های هوش مصنوعی در راستای اصول اخلاقی عمل می‌کنند، ضروری است. این چارچوب‌ها باید به مسائلی مانند انصاف، شفافیت، مسئولیت‌پذیری، حریم خصوصی و کاهش سوگیری بپردازند.

ارزیابی‌های تأثیر اخلاقی: انجام ارزیابی‌های تأثیر اخلاقی قبل از استقرار سیستم‌های هوش مصنوعی می‌تواند به شناسایی خطرات و پیامدهای اخلاقی بالقوه کمک کند. این فرآیند شامل ارزیابی چگونگی تأثیر تصمیمات هوش مصنوعی بر ذینفعان مختلف و اطمینان از اینکه ملاحظات اخلاقی در مراحل طراحی و پیاده‌سازی ادغام شده‌اند، می‌شود.

کمیته‌های اخلاقی: ایجاد کمیته‌های اخلاقی یا هیئت‌های مشورتی در سازمان‌ها می‌تواند تصمیم‌گیری اخلاقی را در پیاده‌سازی هوش مصنوعی هدایت کند. این کمیته‌ها می‌توانند پروژه‌های هوش مصنوعی را بررسی کنند، پیامدهای اخلاقی را ارزیابی کنند و توصیه‌هایی برای رفع نگرانی‌های اخلاقی ارائه دهند.

شفافیت و توضیح‌پذیری

شفافیت: تضمین شفافیت در فرآیندهای تصمیم‌گیری هوش مصنوعی برای ایجاد اعتماد و مسئولیت‌پذیری ضروری است. سازمان‌ها باید در مورد چگونگی تصمیم‌گیری سیستم‌های هوش مصنوعی، داده‌هایی که استفاده می‌کنند و سوگیری‌ها یا محدودیت‌های احتمالی در الگوریتم‌هایشان شفاف باشند.

توضیح‌پذیری: سیستم‌های هوش مصنوعی باید به گونه‌ای طراحی شوند که توضیحات واضح و قابل درکی برای تصمیمات خود ارائه دهند. هوش مصنوعی توضیح‌پذیر به کاربران، ذینفعان و تنظیم‌کننده‌ها کمک می‌کند تا استدلال پشت تصمیمات هوش مصنوعی را درک کنند و هرگونه سوگیری یا خطا را تشخیص دهند.

¹ Balasubramanian & et all

نظارت و ارزیابی مداوم

ممیزی‌های اخلاقی: انجام ممیزی‌های اخلاقی منظم از سیستم‌های هوش مصنوعی می‌تواند به شناسایی و رفع مسائل اخلاقی که ممکن است در طول بهره‌برداری ایجاد شوند، کمک کند. این ممیزی‌ها شامل ارزیابی تأثیر تصمیمات هوش مصنوعی بر ذینفعان، ارزیابی انطباق با دستورالعمل‌های اخلاقی و اعمال تنظیمات برای اطمینان از رعایت استانداردهای اخلاقی است.

سازوکارهای بازخورد: پیاده‌سازی سازوکارهای بازخوردی که به کاربران و ذینفعان اجازه می‌دهد تا در تصمیمات هوش مصنوعی اطلاعاتی ارائه کنند، می‌تواند به سازمان‌ها در بهبود عملکرد اخلاقی سیستم‌های هوش مصنوعی در طول زمان کمک کند. حلقه‌های بازخورد امکان نظارت و ارزیابی مداوم رفتار هوش مصنوعی از منظر اخلاقی را فراهم می‌کنند.

حاکمیت و چارچوب نظارتی هوش مصنوعی

همزمان با تداوم تسریع پذیرش هوش مصنوعی (AI) در بخش مالی، نیاز به سازوکارهای حاکمیتی قوی و چارچوب‌های نظارتی به‌طور فزاینده‌ای اهمیت پیدا می‌کند. پیچیدگی و ابهام سیستم‌های هوش مصنوعی، همراه با پتانسیل آن‌ها برای تأثیرگذاری قابل توجه بر بازارهای مالی و رفاه مصرف‌کننده، مستلزم یک رویکرد جامع برای تضمین توسعه و استقرار اخلاقی و مسئولانه هوش مصنوعی در امور مالی است.

❖ ایجاد هیئت‌ها یا کمیته‌های اخلاق هوش مصنوعی

یک گام حیاتی در ترویج شیوه‌های اخلاقی هوش مصنوعی در مؤسسات مالی، ایجاد هیئت‌ها یا کمیته‌های اخلاق هوش مصنوعی است. این نهادهای تخصصی می‌توانند به عنوان نهادهای نظارتی اختصاصی عمل کنند که مسئول بررسی پروژه‌های هوش مصنوعی، ارزیابی پیامدهای اخلاقی و ارائه راهنمایی در مورد همسویی توسعه و استقرار هوش مصنوعی با اصول اخلاقی هستند (فلوریدی و کالز، ۲۰۱۹). با مشارکت دادن طیف متنوعی از ذینفعان، از جمله متخصصان اخلاق، کارشناسان حقوقی و نمایندگان جامعه، این هیئت‌ها می‌توانند اطمینان حاصل کنند که ملاحظات اخلاقی در فرآیندهای تصمیم‌گیری پیرامون هوش مصنوعی در امور مالی ادغام شده‌اند.

❖ تدوین دستورالعمل‌ها و استانداردهای خاص صنعت

با توجه به چالش‌ها و پیچیدگی‌های منحصربه‌فرد بخش مالی، تدوین دستورالعمل‌ها و استانداردهای خاص صنعت برای استفاده از هوش مصنوعی در امور مالی ضروری است. این دستورالعمل‌ها باید به مسائل اخلاقی کلیدی مانند انصاف، شفافیت، مسئولیت‌پذیری و حریم خصوصی بپردازند، در حالی که چشم‌انداز نظارتی و واقعیت‌های عملیاتی این بخش را نیز در نظر بگیرند (براندریچ و همکاران، ۲۰۱۸). همکاری بین مؤسسات مالی، نهادهای نظارتی و ذینفعان مرتبط می‌تواند ایجاد دستورالعمل‌های جامع و پذیرفته‌شده گسترده را تسهیل کند، سازگاری را تقویت کرده و بهترین شیوه‌ها را در سراسر صنعت ترویج کند.

❖ نقش نهادهای نظارتی در نظارت بر پذیرش هوش مصنوعی

نهادهای نظارتی نقش مهمی در نظارت بر پذیرش هوش مصنوعی در امور مالی و کاهش خطرات مرتبط دارند. این نهادها می‌توانند چارچوب‌هایی را برای حاکمیت و نظارت بر سیستم‌های هوش مصنوعی ایجاد کنند و از انطباق با قوانین و مقررات مربوطه اطمینان حاصل کنند. علاوه بر این، نهادهای نظارتی می‌توانند الزامات شفافیت و افشا را اجباری کنند و به ذینفعان امکان دهند تا فرآیندهای تصمیم‌گیری سیستم‌های هوش مصنوعی را درک کرده و مؤسسات مالی را در قبال اقدامات خود مسئول بدانند (هیئت ثبات مالی، ۲۰۲۰). نهادهای نظارتی می‌توانند با ذینفعان صنعت و کارشناسان

هوش مصنوعی همکاری کنند تا دستورالعمل‌ها و استانداردهایی را برای استفاده مسئولانه از هوش مصنوعی در امور مالی تدوین کنند. این رویکرد مشارکتی می‌تواند به ایجاد تعادل بین ترویج نوآوری و کاهش خطرات احتمالی کمک کند، در حالی که اطمینان حاصل شود که ملاحظات اخلاقی در توسعه و استقرار سیستم‌های هوش مصنوعی گنجانده شده‌اند. با پیاده‌سازی سازوکارهای حاکمیتی قوی و چارچوب‌های نظارتی، بخش مالی می‌تواند از پتانسیل تحول آفرین هوش مصنوعی بهره‌مند شود و در عین حال اصول اخلاقی را در اولویت قرار دهد و از منافع همه ذینفعان محافظت کند. این رویکرد فعال برای حفظ اعتماد عمومی، تضمین شیوه‌های منصفانه و عادلانه و تقویت یک اکوسیستم مالی پایدار و مسئولانه ضروری است.

مطالعات موردی و بهترین روش‌ها

به منظور تشریح بیشتر پیامدهای عملی و راه‌حل‌های بالقوه برای پرداختن به چالش‌های اخلاقی در پذیرش هوش مصنوعی در تصمیم‌گیری‌های مالی، بررسی مطالعات موردی واقعی و نمونه‌هایی از بهترین روش‌ها ارزشمند است. با تجزیه و تحلیل پیاده‌سازی‌های موفق و مواردی که در آن نقض‌های اخلاقی یا بحث‌برانگیزی رخ داده است، ذینفعان می‌توانند بینش‌های ارزشمندی کسب کرده و درک عمیق‌تری از پیچیدگی‌های موجود پیدا کنند.

❖ برجسته کردن چارچوب‌های اخلاقی موفق هوش مصنوعی

چندین مؤسسه و سازمان مالی به طور فعال چارچوب‌های اخلاقی هوش مصنوعی را برای هدایت توسعه و استقرار سیستم‌های هوش مصنوعی در فرآیندهای تصمیم‌گیری خود پیاده‌سازی کرده‌اند. برای مثال، بانک هلندی ING یک شورای مشورتی اخلاق هوش مصنوعی متشکل از کارشناسان داخلی و خارجی تأسیس کرده است تا راهنمایی و نظارت بر استفاده اخلاقی از هوش مصنوعی ارائه دهد (کرافت و همکاران، ۲۰۲۲). این شورا پروژه‌های هوش مصنوعی را بررسی می‌کند، خطرات بالقوه و پیامدهای اخلاقی را ارزیابی می‌کند و اطمینان حاصل می‌کند که سیستم‌های هوش مصنوعی با اصول و ارزش‌های اخلاقی بانک مطابقت دارند.

نمونه دیگر، بانک تجارت امپریال کانادا (CIBC) است که یک منشور اخلاق هوش مصنوعی را توسعه داده و یک شورای اخلاق هوش مصنوعی برای نظارت بر اجرای شیوه‌های اخلاقی هوش مصنوعی ایجاد کرده است (رابرتسون و همکاران، ۲۰۲۱). این شورا شامل نمایندگانی از بخش‌های مختلف از جمله مدیریت ریسک، حقوقی و انطباق است و از وجود یک رویکرد جامع برای رسیدگی به نگرانی‌های اخلاقی اطمینان حاصل می‌کند.

❖ درس‌های آموخته شده از نقض‌های اخلاقی و بحث‌برانگیز

در حالی که نمونه‌هایی از پیاده‌سازی‌های موفق اخلاقی هوش مصنوعی وجود دارد، به همان اندازه مهم است که مواردی را که در آن نقض‌های اخلاقی یا بحث‌برانگیزی رخ داده است، بررسی کنیم. در سال ۲۰۱۹، کارت اعتباری اپل با اتهاماتی مبنی بر تبعیض جنسیتی روبرو شد، با گزارش‌هایی مبنی بر اینکه الگوریتم مورد استفاده برای تعیین سقف اعتباری علیه زنان مغرضانه عمل می‌کند (زندى و همکاران^۱، ۲۰۲۱). این حادثه بر اهمیت رسیدگی به سوگیری‌های احتمالی در سیستم‌های هوش مصنوعی و نیاز به آزمایش و حسابرسی دقیق برای شناسایی و کاهش چنین مسائلی تأکید کرد.

مورد قابل توجه دیگر شامل استفاده از فناوری تشخیص چهره مبتنی بر هوش مصنوعی توسط مؤسسات مالی برای اهداف تأیید هویت بود. این عمل نگرانی‌هایی را در مورد حریم خصوصی و تبعیض احتمالی ایجاد کرد، زیرا نشان داده شده است که الگوریتم‌های تشخیص چهره سوگیری‌هایی را بر اساس نژاد و جنسیت نشان می‌دهند (برانديج و همکاران، ۲۰۱۸).

¹ Zandi & et all

این حوادث بر نیاز به چارچوب‌های حاکمیتی قوی، نظارت مستمر و ادغام ملاحظات اخلاقی در کل چرخه عمر هوش مصنوعی تأکید می‌کنند.

❖ بهترین روش‌ها برای ترویج شفافیت، مسئولیت‌پذیری و انصاف

برای مقابله با چالش‌های اخلاقی و ترویج پذیرش مسئولانه هوش مصنوعی در تصمیم‌گیری‌های مالی، چندین روش برتر ظهور کرده‌اند. این موارد شامل:

- ✓ پیاده‌سازی تکنیک‌های هوش مصنوعی توضیح‌پذیر (XAI) برای افزایش شفافیت و قابلیت تفسیر مدل‌های هوش مصنوعی، و امکان درک فرآیندهای تصمیم‌گیری و شناسایی سوگیری‌های احتمالی برای ذینفعان.
 - ✓ انجام ممیزی‌های منظم الگوریتمی و تست سوگیری برای شناسایی و کاهش پیامدهای تبعیض‌آمیز احتمالی و رفتار ناعادلانه با گروه‌های خاص.
 - ✓ ایجاد تیم‌های متنوع و فراگیر در توسعه و استقرار سیستم‌های هوش مصنوعی، با گنجاندن دیدگاه‌ها و تجربیات متنوع برای کاهش خطر تداوم سوگیری‌ها (فلوریدی و کالز، ۲۰۱۹).
 - ✓ ترویج آموزش و آگاهی مصرف‌کننده در مورد استفاده از هوش مصنوعی در تصمیم‌گیری‌های مالی، توانمندسازی افراد برای انتخاب‌های آگاهانه و مسئولیت‌پذیر کردن مؤسسات.
- با استفاده از این روش‌های برتر و یادگیری از هر دو پیاده‌سازی موفق و نقض‌های اخلاقی، بخش مالی می‌تواند شیوه‌های پذیرش اخلاقی هوش مصنوعی خود را ارتقا دهد و اعتماد، انصاف و مسئولیت‌پذیری را در فرآیندهای تصمیم‌گیری مبتنی بر هوش مصنوعی تقویت کند.

نتیجه‌گیری

در خاتمه، ادغام هوش مصنوعی (AI) در اقتصاد هم‌چالش‌های اخلاقی و هم‌پتانسیل عظیمی برای نوآوری و کارآیی ارائه می‌دهد. چالش‌های اخلاقی استفاده از هوش مصنوعی در امور مالی شامل سوگیری، تبعیض، عدم شفافیت، مسئولیت‌پذیری، ریسک‌های احتمالی برای ثبات بازار و غیره است. این چالش‌ها بر اهمیت نیاز به اتخاذ چارچوب‌ها و دستورالعمل‌های اخلاقی قوی برای اطمینان از استفاده مسئولانه و اخلاقی از هوش مصنوعی در تصمیمات سرمایه‌گذاری تأکید می‌کنند. علی‌رغم چالش‌ها، هوش مصنوعی این پتانسیل را دارد که با بهبود تصمیم‌گیری، بهبود تجربه مشتری و افزایش کارآیی عملیاتی، صنعت مالی را متحول کند. هوش مصنوعی در امور مالی می‌تواند منجر به مدیریت ریسک بهتر و فرآیندهای ساده‌تر شود که در نهایت به نفع مؤسسات مالی و مصرف‌کنندگان است.

برای عبور از چالش‌های اخلاقی و به حداکثر رساندن پتانسیل هوش مصنوعی در اقتصاد، درخواست واضحی برای توسعه دستورالعمل‌های قوی وجود دارد که پذیرش و اتخاذ نقش‌های هوش مصنوعی اخلاقی را ترویج می‌کند. یک رویکرد متعادل برای پذیرش هوش مصنوعی مستلزم ادغام ملاحظات اخلاقی در هر مرحله از پیاده‌سازی هوش مصنوعی، از طراحی و توسعه تا اجرا است. با اولویت قرار دادن استانداردهای اخلاقی، شفافیت و اطمینان از مشارکت انسانی در تصمیم‌گیری‌های هوش مصنوعی، صنعت مالی می‌تواند از پتانسیل کامل هوش مصنوعی استفاده کند و در عین حال خطرات اخلاقی را به حداقل برساند و از منافع ذینفعان محافظت کند. تنظیم‌کننده‌ها، ذینفعان صنعت و تلاش‌های مشترک بین هوش مصنوعی و توسعه‌دهندگان ضروری هستند.

منابع

- ✓ Balasubramanian, S., Devarajan, H. R., Raparathi, M., Dodda, S., Maruthi, S., & Adnyana, I. M. D. (2024). Ethical Considerations in AI-assisted Decision Making for End-of-Life Care in Healthcare. *Power System Technology*, 47, 167-184.
- ✓ Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., ... Amodei, D. (2018). The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. *ArXiv preprint arXiv:1802.07228*. Retrieved from <https://arxiv.org/abs/1802.07228>.
- ✓ Castelnovo, A. (2024). Towards Responsible AI in Banking: Addressing Bias for Fair Decision-Making.
- ✓ Efijemue, O., Ejimofor, I., & Owolabi, O. (2023). Insider Threat Prevention in the US Banking System. *International Journal on Soft Computing*. <https://doi.org/10.5121/ijsc.2023.14302>
- ✓ Fabrègue, B. F. G., & Bogoni, A. (2023). Privacy and Security Concerns in the Smart City. *Smart Cities*, 6(1), 586-613. <https://doi.org/10.3390/smartcities6010027>.
- ✓ Financial Stability Board. (2020). Artificial intelligence and machine learning in financial services. Retrieved from <https://www.fsb.org/wp-content/uploads/P011117.pdf>.
- ✓ Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.8cd550d1>.
- ✓ Jiménez, J. I. (2023). Wine NFT cryptoassets: EU regulatory issues and market challenges. *BIO Web Conf.*, 68, 03026, 1-10. <https://doi.org/10.1051/bioconf/20236803026>.
- ✓ Kirilenko, A., & Lo, A. W. (2013). Moore's law versus Murphy's law: Algorithmic trading and its discontents. *Journal of Economic Perspectives*, 27(2), 51-72. <https://doi.org/10.1257/jep.27.2.51>
- ✓ Kirilenko, A., Kyle, A. S., Samadi, M., & Tuzun, T. (2017). The flash crash: High-frequency trading in an electronic market. *The Journal of Finance*, 72(3), 967-998.
- ✓ Mandych, O., Staverska, T., & Maliy, O. (2023). Integration of Artificial Intelligence into the Blockchain and Cryptocurrency Market. <https://doi.org/10.31891/mdes/2023-10-8>.
- ✓ OECD. (2023). Generative artificial intelligence in finance. *OECD Artificial Intelligence*.
- ✓ Olatoye, F. O., Awonuga, K. F., Mhlongo, N. Z., Ibeh, C. V., Elufioye, O. A., & Ndubuisi, N. L. (2024). AI and ethics in business: A comprehensive review of responsible AI practices and corporate responsibility.
- ✓ Pavashe, A. S., Kadam, P. D., Zirange, V. B., & Katkar, R. D. (2023). The Impact of Artificial Intelligence on Employment and Workforce Trends in the Post-Pandemic Era.
- ✓ Saikanth, D. R. K. (2024). The Impact of Emerging Technologies on Sustainable Agriculture and Rural Development. *IJECC*, 14(1), 253-263.
- ✓ Schwendicke, F., & Krois, J. (2021). Better Reporting of Studies on Artificial Intelligence: CONSORT-AI and Beyond. *Journal of Dental Research*, 100(7), 677-680.