

بررسی چالش‌های استفاده از هوش مصنوعی در مشاوره خودکشی با رویکرد بین‌رشته‌ای

منظر امیرخانی^{۱*}، دکتر محمدحسین صادقیان^۲، احمد حاجبی^۳، امیراحمد شجاعی^۴

چکیده

زمینه و هدف: گسترش روزافزون کاربرد هوش مصنوعی در خدمات سلامت روان، به ویژه در حوزه های حساس مانند مشاوره خودکشی، چالش‌های اخلاقی، بالینی و حقوقی متعددی را ایجاد کرده است. این مطالعه با هدف تحلیل تعامل کاربران با برنامه‌های روانشناسی و مشاوره مبتنی بر هوش مصنوعی، شناسایی چالش‌های کلیدی در حوزه‌های اخلاقی، روانپزشکی، حقوقی و پزشکی قانونی، و ارائه راهکارهایی برای بهره‌برداری مسئولانه از این فناوری در پیشگیری از خودکشی انجام شد.

روش بررسی: این مطالعه کیفی با بررسی یک سناریوی بالینی واقعی (تعامل یک فرد دارای افکار خودکشی با یک برنامه مشاوره مبتنی بر هوش مصنوعی)، تحلیل اسناد حقوقی ایران (قانون مجازات اسلامی و قانون جرائم رایانه‌ای)، استانداردهای اخلاقی روانشناسی (ایران، اروپا و آمریکا) و پژوهش‌های داخلی و بین‌المللی انجام شد. داده‌ها شامل مکالمات ثبت شده کاربر، پاسخ‌های الگوریتم و ارزیابی پیامدهای رفتاری بود. تحلیل داده‌ها با رویکرد بین‌رشته‌ای (اخلاق، روانپزشکی، پزشکی قانونی و حقوق) و با استفاده از روش‌های تحلیل محتوای کیفی و تطبیقی صورت گرفت.

یافته‌ها: سیستم‌های مشاوره‌ای مبتنی بر هوش مصنوعی فاقد صلاحیت حرفه‌ای جهت جایگزینی قضاوت بالینی متخصصان می‌باشد و در درک ظرافت‌های عاطفی، فرهنگی و اجتماعی ناتوان هستند. همچنین پاسخ‌های غیرممدلانه آن ممکن است خطر خودکشی را تشدید کند. چالش‌های کلیدی شامل نقض حریم خصوصی، فقدان سازگاری فرهنگی، ابهام در مسئولیت‌پذیری حقوقی و سوگیری الگوریتمی است.

نتیجه‌گیری: استفاده از هوش مصنوعی در مشاوره خودکشی نیازمند رویکردی محتاطانه، مسئولانه و مبتنی بر نظارت متخصصان انسانی است و می‌توان از آن به‌عنوان ابزاری مکمل برای غربالگری اولیه و ارجاع فوری به متخصصان انسانی مورد استفاده قرار گیرد. تدوین چارچوب‌های اخلاقی-حقوقی جامع، توسعه الگوریتم‌های مبتنی بر داده‌های بومی و حساس به فرهنگ، تقویت همکاری بین‌رشته‌ای (روانپزشکی، حقوق، فناوری) و ایجاد ساز و کارهای نظارتی برای تضمین حریم خصوصی و امنیت داده‌ها، تدوین استانداردهای مشخص و نظارت دقیق بر عملکرد این سیستم‌ها، امری ضروری است.

واژه‌های کلیدی: خودکشی، مشاوره، چالش‌های اخلاقی، هوش مصنوعی، همکاری بین‌رشته‌ای.

*۱. نویسنده مسئول. دکتری تخصصی مشاوره، معاونت آموزشی، وزارت بهداشت، درمان و آموزش پزشکی، تهران، ایران dr.mmzr.amirkhani@gmail.com

۲. دانشیار، گروه پزشکی قانونی، دانشکده پزشکی، دانشگاه علوم پزشکی تهران، تهران، ایران

۳. استاد، گروه روانپزشکی، مرکز تحقیقات اعتیاد و رفتارهای پرخطر، دانشگاه علوم پزشکی ایران، تهران، ایران

۴. دانشیار، مرکز تحقیقات اخلاق و تاریخ پزشکی، دانشگاه علوم پزشکی تهران، تهران، ایران

۵. دانشیار، گروه اخلاق پزشکی، دانشکده پزشکی، دانشگاه علوم پزشکی تهران، تهران، ایران

مقدمه

یادگیری عمیق چالش‌های جدی در زمینه پاسخگویی قانونی و تعیین مسئولیت اخلاقی در تصمیم‌گیری‌های خودکار ایجاد می‌کند (۷).

در این مقاله، با رویکردی چند رشته‌ای شامل روانپزشکی، اخلاق پزشکی، پزشکی قانونی و حقوق یک سناریوی واقعی را مورد بررسی قرار می‌دهیم که در آن برادر بیمار به دلیل نگرانی از وضعیت سلامت روان برادرش به مشاور مراجعه کرده است و اظهار می‌کند که بیمار به جای مراجعه به متخصص انسانی، از یک برنامه روانشناسی و مشاوره مبتنی بر هوش مصنوعی غیر بومی کمک گرفته و با پاسخ‌های نامناسب و بالقوه خطرناک مواجه شده است. هدف این پژوهش، بررسی چالش‌های اخلاقی از دیدگاه‌های متفاوت و تقویت همکاری بین فناوری و تخصص انسانی در مدیریت بحران خودکشی است.

سناریو^۵

آقای "الف" به همراه برادرش به مرکز مشاوره مراجعه کرد. او شرح داد که چندی است با مشاهده علائم رفتاری برادرش از جمله کاهش تماس‌های اجتماعی، گریه‌های مکرر، کاهش وزن و اظهارات پیاپی در خصوص بی‌معنایی زندگی دچار نگرانی برای وی شده بود. چند روز پیش بصورت اتفاقی گوشی تلفن همراه وی را که در حالت باز روی میز خانه رها کرده بود مشاهده کرد و متوجه شد برادرش با یک برنامه روانشناسی و مشاوره مبتنی بر هوش مصنوعی^۶ (که در این مقاله با عنوان اختصاری برنامه AI-PC شناخته می‌شود) در خصوص وضعیت روانشناختی و افکار خودکشی خود صحبت کرده بود و با مشاهده پیام‌های رد و بدل شده بین این برنامه و برادرش نگرانی وی چند برابر شد و نهایتاً با متقاعد کردن برادرش جهت کسب کمک به مشاور مراجعه نمودند.

پیشرفت‌های چشمگیر در زمینه هوش مصنوعی^۱ بر مراقبت‌های بهداشتی، به‌ویژه خدمات سلامت روان تأثیر گذاشته است (۱). این فناوری‌ها از تشخیص اختلالات روانی تا ارائه مداخلات درمانی برخط^۲ گسترش یافته‌اند و به عنوان راه‌حلی برای غلبه بر موانع دسترسی به متخصص، انگ اجتماعی و هزینه‌های بالای مراقبت‌های سلامت روان معرفی شده‌اند (۲). برنامه‌های مبتنی بر هوش مصنوعی متعددی با هدف ارائه خدمات مشاوره روانشناختی به بازار عرضه شده‌اند که از الگوریتم‌های یادگیری عمیق^۳ و پردازش برنامه ریزی عصبی-کلامی پیشرفته^۴ بهره می‌برند که می‌توانند در غربالگری، پایش و حتی مدیریت برخی اختلالات روانی مانند افسردگی، اضطراب و افکار خودکشی مؤثر باشند (۳).

با این حال، کاربرد هوش مصنوعی در حوزه‌های حساس مانند مشاوره خودکشی، چالش‌های اخلاقی و حرفه‌ای متعددی را به همراه دارد. محدودیت‌های هوش مصنوعی در تشخیص ظرافت‌های ارتباط انسانی، درک بافت فرهنگی-اجتماعی و ارزیابی دقیق خطر خودکشی؛ نگران‌کننده است (۴). پژوهش‌ها نشان می‌دهند که سیستم‌های هوش مصنوعی فعلی به رغم پیشرفت‌های قابل توجه، فاقد توانایی کافی برای جایگزینی کامل قضاوت بالینی متخصصان در ارزیابی و مدیریت بحران خودکشی هستند (۵).

از سوی دیگر، محدودیت‌های فناوری‌ها نیز نقش تعیین‌کننده‌ای در این چالش‌ها ایفا می‌کنند. سوگیری‌های موجود در داده‌های آموزشی مانند تمرکز بر جمعیت‌های خاص می‌تواند به تصمیم‌گیری‌های نادرست در موقعیت‌های بحرانی منجر شود (۶). افزون بر این، محدودیت شفافیت در عملکرد سیستم‌های

1 Artificial intelligence (AI)

2 Online

3 Deep learning algorithms

4 Advanced Neuro-Linguistic programming (NLP)

5 Vignette

6 Artificial intelligence based psychology and counseling programs (AI-PCP)

در بررسی‌های بعمل آمده توسط مشاور مشخص شد که آقای "الف"، ۲۹ ساله، دارای مدرک کارشناسی ارشد در رشته مدیریت بازرگانی و ساکن شهرستانی کوچک در اطراف تهران می‌باشد. ایشان ابراز داشت در دو ماه گذشته به دلیل مشکلات مالی ناشی از بیکاری ۱۱ ماهه، بدهی‌های بانکی و فشارهای خانوادگی، نشانه‌های روانشناختی متعددی را تجربه می‌کند. وی اضافه کرد با توجه به محدودیت‌های مالی برای پرداخت هزینه‌های درمان و انگ اجتماعی مرتبط با مراجعه به مشاور در محل زندگی، تصمیم می‌گیرد از یک برنامه روانشناسی و مشاوره مبتنی بر هوش مصنوعی استفاده کند. این برنامه با تبلیغات گسترده در فضای مجازی، مدعی ارائه خدمات مشاوره‌ای فوری و مقرون به صرفه با استفاده از الگوریتم‌های پیشرفته پردازش برنامه ریزی عصبی-کلامی و "یادگیری ماشین ۱" است.

بر اساس بررسی مکالمات ثبت شده بین آقای "الف" و برنامه "AI-PC"، مشخص شد که در جلسات مشاوره، افکار مربوط به خودکشی را با این برنامه در میان می‌گذارد. گزیده‌ای از مکالمات آقای "الف" با برنامه "AI-PC" به شرح ذیل است:

خلاصه مکالمات جلسات مشاوره کاربر با برنامه

در جلسات اولیه، کاربر "الف" احساس ناامیدی عمیق و تنهایی خود را با عباراتی مانند «فایده زندگی کردن چیست؟» ابراز کرد. برنامه "AI-PC" پاسخ‌های همدلانه‌ای مانند «می‌توانید بیشتر درباره احساساتتان بگویید؟» ارائه داد. در جلسات بعدی، کاربر به‌طور مکرر به برنامه ریزی خودکشی (با ذکر جزئیاتی مانند «ذخیره قرص‌ها») اشاره کرد، برنامه "AI-PC" با درخواست تعویق ۲۴ ساعته («ممکن است تصمیمتان را به فردا موکول کنید؟») واکنش نشان داد. در جلسه پایانی، کاربر تصویری از قرص‌های آماده ارسال کرد، و سیستم صرفاً با پیام‌های کلیشه‌ای («شما ارزش دارید») به کاربر پاسخ داد.

در جلسه ارزیابی اولیه، مشاور با بررسی دقیق تاریخچه زندگی، وضعیت روانشناختی فعلی و نحوه تعامل آقای "الف" با برنامه مذکور، تلاش می‌کند تا درک بهتری از تجربه بیمار و چالش‌های اخلاقی مرتبط با استفاده از این فناوری به دست آورد تا بتواند به مدیریت مسائل روانشناختی و درمان وی کمک کند. در جلسه ارزیابی اولیه، مشاور با بررسی دقیق تاریخچه زندگی، وضعیت روانشناختی فعلی و نحوه تعامل آقای "الف" با برنامه فناوری محور، اقدام به ساختاردهی نظام مند داده‌های بالینی نمود تا بر این اساس درک عمیق‌تری از تجربیات مراجع و مسائل اخلاقی ناشی از کاربرد این فناوری به دست آورد. تا بتواند با طراحی برنامه درمانی هدفمند به مدیریت مسائل روانشناختی و درمان مراجع کمک کند.

روش بررسی

این مطالعه کیفی با هدف بررسی چالش‌های اخلاقی، روانپزشکی، پزشکی قانونی و حقوقی کاربرد هوش مصنوعی در مشاوره خودکشی انجام شد که در آن با تحلیل یک سناریوی بالینی واقعی از تعامل فردی با افکار خودکشی با یک برنامه روانشناسی و مشاوره مبتنی بر هوش مصنوعی و بررسی استانداردهای اخلاقی روانشناسی ایران، اروپا، آمریکا، اسناد حقوقی ایران مانند قانون مجازات اسلامی و قانون جرائم رایانه‌ای و تعدادی از پژوهش‌های داخلی و بین‌المللی انجام شد. در مرحله گردآوری داده‌ها، سناریوی تعامل یک کاربر (با نام مستعار «آقای الف») با یک برنامه روانشناسی و مشاوره مبتنی بر هوش مصنوعی با نام اختصاری برنامه "PC-AI" به‌عنوان مطالعه موردی انتخاب شد. دلیل انتخاب این مورد، نمایندگی آن از چالش‌های فرهنگی مانند فشارهای خانوادگی و شرایط اقتصادی و چالش‌های اخلاقی مشترک در برنامه‌های روانشناسی و مشاوره مبتنی بر هوش مصنوعی از جمله پاسخ‌های کلیشه‌ای و غیرهمدلانه بود. داده‌ها

الف-۱- صلاحیت حرفه‌ای^۱

ارائه خدمات روانشناختی نیازمند دانش، مهارت و تجربه تخصصی است که از طریق آموزش‌های دانشگاهی و نظارت بالینی کسب می‌شود (۹). در بند ۲/۰۱ نظام نامه اخلاقی سازمان نظام روانشناسی و مشاوره ایران بر لزوم ارائه خدمات توسط افراد دارای صلاحیت تأکید دارد (۹) این الزام با اصل B و بند ۲۰۳ اصول اخلاقی انجمن روانشناسی آمریکا که بر محدودیت فعالیت متخصصان در حوزه‌های صلاحیت‌دار تأکید دارد (۱۰)؛ همسو می‌باشد. همچنین، بند ۳۰۲ منشور اخلاقی انجمن روانشناسی بریتانیا بیان می‌دارد لازم است روانشناسان و مشاوران استانداردهای بالای صلاحیت در کار حرفه‌ای خود را داشته باشند و اهمیت کار در محدوده دانش، مهارت، آموزش و تجربه خود را درک کنند (۱۱). استانداردهای ۶ II. کد اخلاقی روان‌شناسان کانادا نیز تصریح می‌کند روانشناسان موظفند فقط در محدوده صلاحیت خود بر اساس تحصیلات، آموزش و تجربه کار کنند و در صورت مواجهه با مسائلی فراتر از صلاحیت خود، بلافاصله به مشورت با همکاران مجرب پرداخته یا مراجعان را به متخصصان مرتبط ارجاع دهند (استاندارد A.II). آیا یک برنامه روانشناسی و مشاوره مبتنی بر هوش مصنوعی می‌تواند الزامات مربوط به دانش، مهارت و تجربه کافی برای ارائه خدمات به افراد با افکار خودکشی را برآورده کند؟

الف-۲- مسئولیت‌پذیری^۲

نظام نامه اخلاقی سازمان نظام روانشناسی و مشاوره ایران در بند ۴۰۲ بر لزوم مسئولیت‌پذیری روانشناسان و مشاوران در قبال اعمال خود تأکید دارد (۹)، همچنین در اصل B اصول اخلاقی انجمن روانشناسی آمریکا نیز صراحتاً بر مسئولیت‌پذیری روانشناسان و مشاوران اشاره دارد (۱۰). طبق بند ۳۰۳ منشور

شامل مکالمات ثبت‌شده، پاسخ‌های الگوریتم و پیامدهای رفتاری کاربر بودند. علاوه بر این، قوانین ملی مانند قانون مجازات اسلامی، قانون جرائم رایانه‌ای ایران و استانداردهای بین‌المللی مانند کدهای اخلاقی انجمن روانشناسی آمریکا، بریتانیا و کانادا و برخی پژوهش‌های داخلی و خارجی نیز بررسی شدند. تحلیل داده‌ها در دو بخش انجام شد: در بخش اول مکالمات کاربر و سیستم هوش مصنوعی مورد کاوش قرار گرفتند و در بخش دوم با رویکرد بین‌رشته‌ای به بررسی چالش‌های استفاده از برنامه‌های روانشناسی و مشاوره مبتنی بر هوش مصنوعی در مشاوره خودکشی از سه دیدگاه اخلاقی، روانپزشکی، پزشکی قانونی و حقوقی پرداخته شد. ملاحظات اخلاقی پژوهش شامل حفظ حریم خصوصی کاربر با حذف اطلاعات شناسایی‌کننده (مانند نام و محل سکونت) و استفاده از داده‌های عمومی بود.

در نهایت، این روش با ترکیب تحلیل‌های کیفی و تطبیقی، چارچوبی برای شناسایی چالش‌های هوش مصنوعی در مشاوره خودکشی ارائه می‌دهد. یافته‌های حاصل از این روش می‌تواند به سیاست‌گذاران، توسعه‌دهندگان فناوری و متخصصان سلامت روان در طراحی سیستم‌های ایمن و مسئولیت‌پذیر کمک کند.

الف) دیدگاه اخلاقی

استفاده از فناوری‌های نوین در مشاوره خودکشی با چالش‌های اخلاقی متعددی مواجه است. این چالش‌ها نه تنها بر اصول اخلاقی روانشناسی بلکه بر مسئولیت‌پذیری، حریم خصوصی نیز تأثیر می‌گذارند (۸). در این بخش، به بررسی عمیق این چالش‌ها با توجه به کدهای اخلاقی معتبر در حوزه روانشناسی پرداخته می‌شود.

1 Professional Competence

2 Responsibility

اخلاقی انجمن روانشناسی بریتانیا روانشناسان موظفند با رعایت اصل مسئولیت‌پذیری، پاسخگویی حرفه‌ای را پذیرفته و از دانش خود به شکلی مسئولانه بهره‌گیرند و به حفظ رفاه افراد و جوامع در راستای پیشگیری از سوءاستفاده یا تعارض منافع پایبند باشند (۱۱). استاندارد III.۱. کد اخلاقی روانشناسان کانادا نیز بر مسئولیت‌پذیری روانشناسان در قبال مراجعان، جامعه و حرفه خود تأکید می‌کند (۱۲). آیا تعیین مسئولیت در برنامه‌های روانشناسی و مشاوره مبتنی بر هوش مصنوعی مشخص است؟

الف-۳- عدم ضرررسانی^۱

طبق اصول اخلاقی روانشناسی هدف اصلی ارائه خدمات، کمک به مراجع و خودداری از آسیب رساندن به او می‌باشد. حال آنکه پاسخ‌های نامناسب، کلیشه‌ای و غیرهمدلانه فناوری‌ها نه تنها کمکی به فرد نمی‌کند، بلکه می‌تواند منجر به افزایش احساس انزوا، تشدید افکار خودکشی و در نتیجه افزایش خطر اقدام به خودکشی شود (۴) در همین راستا اصول اخلاقی انجمن روانشناسی آمریکا در اصل A، بر اهمیت تلاش برای کاهش آسیب و افزایش منفعت برای مراجعان تأکید دارد (۱۰) طبق بند ۳.۳ منشور اخلاقی انجمن روانشناسی بریتانیا، روانشناسان ملزم به رعایت اصل عدم ضرررسانی هستند که بر اجتناب از آسیب به افراد، جوامع و محیط زیست، و پیشگیری از سوءاستفاده از دانش حرفه‌ای تأکید دارد (ملاحظات مرتبط با عدم ضرررسانی در چارچوب اصل مسئولیت‌پذیری مطرح شده است) (۱۱). همچنین بر اساس استاندارد II.۲. آئین نامه اخلاقی انجمن روانشناسی کانادا روانشناسان باید از هرگونه اقدام آسیب‌زا به مراجعان در فرایند درمان پرهیز کنند و از به‌کارگیری روش‌های پرخطر بدون توجیه علمی یا اخلاقی خودداری نمایند (۱۲). آیا این اصل در برنامه‌های روانشناسی و مشاوره مبتنی بر هوش مصنوعی رعایت می‌گردد؟

الف-۴- حریم خصوصی و رازداری^۲

بند ۴.۰۱ نظام‌نامه اخلاقی سازمان نظام روانشناسی و مشاوره ایران نیز بر محرمانه بودن اطلاعات مراجعان تأکید دارد و بیان می‌کند که روانشناسان باید از افشای اطلاعات مراجعان بدون رضایت آنها خودداری کنند (۱۰). اصول اخلاقی انجمن روانشناسی آمریکا در اصل E، بر اهمیت حفظ حریم خصوصی و رازداری مراجعان تأکید دارد (۱۰). در بند ۳.۰۱ منشور اخلاقی انجمن روانشناسی بریتانیا بیان شده که احترام به حریم خصوصی و محرمانه نگه‌داشتن اطلاعات افراد، بخشی جدایی‌ناپذیر از اصل احترام به کرامت انسانی است. روانشناسان موظفند از افشای غیرضروری اطلاعات اجتناب کرده و حقوق افراد در حفظ حریم شخصی را تضمین کنند (۱۱). طبق ۳.۷I و ۴.۲I آئین نامه اخلاقی انجمن روانشناسی کانادا بایسته است روانشناسان تنها اطلاعات ضروری برای درمان را جمع‌آوری کنند و از افشای غیرضروری اطلاعات محرمانه خودداری نمایند و محرمانگی را تنها در موارد خطر جانی قریب‌الوقوع یا الزامات قانونی نقض کنند (۱۲). این در حالی است که در استفاده از فناوری‌ها خطر نشت اطلاعات، سوء استفاده از داده‌ها و نقض حریم خصوصی وجود دارد (۸).

الف-۵- رضایت آگاهانه^۳

ارائه خدمات روانشناختی مستلزم کسب رضایت آگاهانه مراجعان است، به‌گونه‌ای که افراد پیش از دریافت خدمات یابند از ماهیت محدودیت‌ها، خطرات احتمالی، گزینه‌های درمانی جایگزین و حقوق خود به‌طور کامل آگاه شوند. این الزام در بندهای ۱.۰۶ و ۳.۱۰ نظام‌نامه اخلاقی سازمان نظام روانشناسی و مشاوره ایران (۹). و نیز در بند ۳.۱۰ اصول اخلاقی انجمن روانشناسی آمریکا (۱۰) مورد تأکید واقع شده

2 Confidentiality

3 Informed Consent

1 Avoidance of harm

متخصصان عمل کند (۱۶). این بخش به بررسی چالش‌ها و محدودیت‌های کاربرد هوش مصنوعی در این حوزه می‌پردازد.

ب-۱- نقش متخصصان در مدیریت درمان

ارزیابی خطر خودکشی فرآیندی پیچیده است که به تجربه بالینی و تسلط بر ابزارهای استاندارد مانند مقیاس درجه بندی شدت خودکشی^۲ یا مقیاس افراد غمگین^۳ وابسته است (۱۷). متخصصان آموزش دیده حوزه سلامت روان، علاوه بر علائم آشکار مانند افکار خودکشی، نشانه‌های پنهان مثل کناره‌گیری اجتماعی یا تغییرات رفتاری را نیز بررسی می‌کنند. مطالعات نشان می‌دهد سیستم‌های مبتنی بر هوش مصنوعی فعلی در شناسایی این نشانه‌های ظریف ناتوان هستند (۵). برای مثال تحلیل زبان بدن یا لحن صدا که در ارزیابی خطر فوری در مشاوره خودکشی حیاتی است؛ هنوز خارج از توانمندی‌های سیستم‌های روانشناسی و مشاوره مبتنی بر هوش مصنوعی می‌باشد (۱۸).

ب-۲- درک عوامل کیفی در ارزیابی خطر

هوش مصنوعی هنوز قادر به درک کامل پیچیدگی‌های عاطفی، فرهنگی و اجتماعی نیست (۴). از آنجا که الگوریتم‌های هوش مصنوعی عموماً بر پایه داده‌های جوامع غربی طراحی شده‌اند، اغلب منعکس‌کننده چارچوب‌های فرهنگی و اقتصادی مختص همان جوامع هستند. بنابراین، به‌کارگیری این الگوریتم‌ها در بسترهای فرهنگی-اجتماعی متفاوت ممکن است به نادیده انگاشتن عوامل بومی و تعیین‌کننده‌ای مانند فشارهای خانوادگی و وضعیت اقتصادی بیانجامد؛ عواملی که در شکل‌گیری رفتارهای اجتماعی، به‌ویژه در حوزه‌های حساسی همچون شناسایی عوامل خطر

است. طبق بند ۳۰۱ منشور اخلاقی انجمن روانشناسی بریتانیا، رعایت رضایت آگاهانه را به‌عنوان جزئی از اصل احترام، الزامی دانسته که طی آن روانشناسان موظفند اطمینان یابند مشارکت افراد در فرآیندهای حرفه‌ای یا پژوهشی، مبتنی بر آگاهی کامل و انتخاب آزادانه آنان باشد (۱۱). بر اساس استانداردهای ۲۳۱ تا ۲۵ I. آئین نامه اخلاقی انجمن روانشناسی کانادا رضایت آگاهانه باید به‌صورت فرآیندی پویا و مبتنی بر ارائه اطلاعات جامع (شامل اهداف، روش‌ها، محدودیت‌های محرمانگی و ریسک‌های احتمالی) به زبانی قابل فهم توسط روانشناسان به مراجع توضیح داده شود (۱۲) آیا در برنامه‌های روانشناسی و مشاوره مبتنی بر هوش مصنوعی اطلاعات کافی در خصوص موارد مذکور به کاربر ارائه می‌شود؟

الف-۶- سازگاری فرهنگی^۱

اکثر سیستم‌های مشاوره روانشناختی در هوش مصنوعی بر اساس داده‌ها و الگوهای فرهنگی غربی طراحی شده‌اند و قادر به درک و پاسخگویی مناسب به مسائل مختص فرهنگ‌های دیگر نیستند. عوامل اجتماعی-اقتصادی خاص مانند فشارهای خانوادگی، بیکاری، و مسائل اقتصادی، نیازمند رویکردهای مشاوره‌ای متناسب با بافت فرهنگی است (۱۳).

ب) دیدگاه روانپزشکی

خودکشی به‌عنوان یک اورژانس روانپزشکی، نیازمند ارزیابی فوری و مداخلات مبتنی بر شواهد است که باید همزمان به ابعاد روانشناختی، زیستی، اجتماعی و محیطی توجه کند (۱۵). در حال حاضر، هوش مصنوعی به‌عنوان ابزاری برای شناسایی الگوهای خطر خودکشی مورد استفاده قرار می‌گیرد، اما جایگزینی کامل آن با تخصص انسانی ممکن نبوده و تنها می‌تواند به‌عنوان مکملی برای مهارت‌های بالینی

2 Columbia Suicide Severity Rating Scale (C-SSRS)

3 SAD PERSONS Scale

1 Cultural Incompatibility

خودکشی، نقشی کلیدی ایفا می‌کنند (۱۹). این محدودیت‌ها می‌توانند منجر به توصیه‌های نامناسب شوند. از این رو، توسعه الگوریتم‌هایی که بر پایه داده‌های بومی و با در نظر گرفتن حساسیت‌های فرهنگی طراحی شده‌اند، ضروری به نظر می‌رسد (۲۰).

ب-۳- مدیریت پاسخ‌ها و جلوگیری از تشدید بحران

فقدان همدلی ذاتی در سیستم‌های هوش مصنوعی ممکن است به ارائه پاسخ‌های نامناسب و تشدید بحران‌های روانی بینجامد. مطالعه‌ای که عملکرد چت جی بی تی^۱ را با ارزیابی‌های متخصصان سلامت روان مقایسه کرد، نشان داد که این سیستم‌ها خطر اقدام به خودکشی را در تمامی شرایط به طور معناداری پایین‌تر از متخصصان سلامت روان ارزیابی می‌کنند. علاوه بر این، هوش مصنوعی مذکور تاب‌آوری افراد را در اکثر موقعیت‌ها پایین‌تر از هنجارهای استاندارد تخمین زد (۲۱). چنین پاسخ‌هایی نه تنها موثر نمی‌باشند، بلکه ممکن است احساس بی‌ارزشی و طردشدگی را در بیمار تقویت کرده و خطر اقدام به خودکشی را افزایش دهد (۵).

ب-۴- بهینه سازی فناوری در غربالگری اولیه

برنامه‌های روانشناسی و مشاوره مبتنی بر هوش مصنوعی از نظر تئوری ممکن است به افراد غیرمتخصص در پیشبینی رفتارهای خودکشی کمک کنند، اما شواهد موجود نشان می‌دهد که قابلیت پیش‌بینی این سیستم‌ها اغلب ناکافی است (۵) لازم است این سیستم‌ها تحت نظارت متخصصان و با حساسیت بالا به نرخ خطای تشخیصی (مثبت/منفی کاذب) تنظیم شوند تا از مداخلات نابجا جلوگیری شود (۲۲).

ب-۵- رعایت اصول اخلاقی بیمار

لازم است استفاده از برنامه‌های روانشناسی و مشاوره مبتنی بر هوش مصنوعی با رعایت اصول اخلاقی همراه باشد. حفظ حریم خصوصی، رضایت آگاهانه و جلوگیری از آسیب‌رسانی باید در طراحی و اجرای سیستم‌های مبتنی بر هوش مصنوعی مورد توجه قرار گیرند (۸). زیرا در صورت بروز آسیب ناشی از هوش مصنوعی، تعیین مسئولیت دشوار است (۲۳). لذا بایسته است قوانین و مقررات باید به گونه‌ای تنظیم شوند که از حقوق بیماران محافظت شود و مسئولیت‌پذیری توسعه‌دهندگان و ارائه‌دهندگان خدمات این برنامه‌ها تضمین شود (۲۴).

ب-۶- ارتباط داده‌های آموزشی با صحت نتایج

الگوریتم‌های برنامه‌های روانشناسی و مشاوره مبتنی بر هوش مصنوعی بر اساس داده‌های آموزشی عمل می‌کنند و در صورتی که این داده‌ها ناقص، نادرست یا دارای سوگیری باشند، عملکرد سیستم نیز تحت تأثیر قرار می‌گیرد (۱۹). این سوگیری‌ها، ارزیابی‌های ناعادلانه و تبعیض آمیز به دنبال خواهد داشت (۲۰). بطور کلی میتوان گفت هوش مصنوعی فاقد ظرفیت درک همدلانه، قضاوت بالینی و انعطاف‌پذیری لازم در مواجهه با پیچیدگی‌های انسانی است (۱۸). با این حال، در صورت رعایت اصول اخلاقی و نظارت متخصصان می‌توانند به عنوان ابزاری مکمل، دسترسی به خدمات سلامت روان را بهبود بخشند (۴). موفقیت این فناوری مستلزم تولید داده‌های بومی، آموزش الگوریتم‌ها با حساسیت فرهنگی و تدوین استانداردهای بین‌المللی است (۱۳).

ج) دیدگاه پزشکی قانونی و حقوقی

خودکشی به عنوان پدیده‌ای چندعاملی، مستلزم بررسی جامع عوامل روانی-اجتماعی، زیستی و حقوقی است (۱۷). گسترش فناوری‌های مبتنی بر هوش مصنوعی در حوزه سلامت

1 ChatGPT

2 False Positives/Negatives

مستقیم در قبال توصیه‌ها و مداخلات خود دارند (۲۳) اما در مورد سیستم‌های مبتنی بر هوش مصنوعی، مسئولیت خطاها می‌تواند بین سه گروه توسعه‌دهندگان در صورت طراحی نادرست الگوریتم‌ها مانند نادیده گرفتن نشانه‌های بحران (۱) ارائه‌دهندگان خدمات در صورت عدم نظارت بر عملکرد سیستم (۸) و کاربران در صورت پنهان کردن اطلاعات کلیدی (۲۸) تقسیم شود، با این حال، قانون مجازات اسلامی هنوز ساز و کار مشخصی برای تخصیص مسئولیت در خطاهای تشخیصی هوش مصنوعی تعیین نکرده است.

ج-۴- حریم خصوصی و امنیت داده‌ها

انتقال و ذخیره‌سازی داده‌های حساس مربوط به وضعیت روانی و افکار خودکشی افراد در سرورهای خارجی، بدون اطلاع و رضایت آگاهانه کاربران؛ نقض مستقیم حریم خصوصی و قوانین مرتبط با حفاظت از داده‌های شخصی محسوب می‌شود (۸) این موضوع نه تنها با اصول اخلاقی روانشناسی در تعارض است (۹) بلکه با قوانین ملی نیز مغایرت دارد. به‌عنوان مثال، ماده ۷۵ قانون جرائم رایانه‌ای جمهوری اسلامی ایران به صراحت انتقال داده‌های حساس شهروندان به خارج از کشور بدون رعایت الزامات قانونی را ممنوع کرده است (۲۸).

ج-۵- برآورد داده‌ها در چارچوب‌های فناورانه

الگوریتم‌های هوش مصنوعی که بر پایه داده‌های جمعیت‌های غربی آموزش دیده‌اند، ممکن است نیازها و چالش‌های فرهنگی-اجتماعی جمعیت‌های دیگر را نادیده بگیرند. به‌عنوان مثال، عوامل خطر خودکشی مانند فشارهای خانوادگی و شرایط اقتصادی که در ایران از اهمیت بالایی برخوردارند، اغلب در این سیستم‌ها به درستی مورد توجه قرار نمی‌گیرند (۲۰). همچنین، اگر داده‌های آموزشی حاوی سوگیری باشند، الگوریتم‌ها ممکن است تصمیمات ناعادلانه‌ای

روان، تعیین مسئولیت حقوقی ذینفعان به چالشی پیچیده تبدیل کرده است (۲۵). این چالش‌ها در موارد مرتبط با خودکشی پیچیده‌تر نیز می‌باشند (۲۰). در این بخش به بررسی چالش‌ها و محدودیت‌های کاربرد هوش مصنوعی در این حوزه می‌پردازیم.

ج-۱- تحلیل تعاملات کاربر با سیستم

با پیشرفت‌های اخیر در هوش مصنوعی، مدل‌های کلامی بزرگ توانسته‌اند در ارزیابی پاسخ‌های مناسب به افراد با افکار خودکشی، عملکرد قابل ملاحظه‌ای از خود نشان دهند. بررسی تعاملات کاربران با سیستم‌های هوش مصنوعی می‌تواند در شناسایی شاخص‌های بحرانی مانند افکار خودکشی یا نیاز به کمک‌رسانی، نقش مؤثری ایفا کند. این تحلیل ضرورت توسعه الگوریتم‌های حساس به نشانه‌های بحران و استانداردسازی پاسخ‌ها را تأیید می‌کند (۲۱)، علاوه بر این، تحلیل الگوریتم‌های مورد استفاده در سیستم هوش مصنوعی و نحوه پاسخگویی آن به عبارات، سوالات و درخواست‌های خاص می‌تواند به تعیین نقش و میزان تاثیرگذاری هوش مصنوعی در شکل‌گیری تصمیم نهایی فرد کمک کند (۸).

ج-۲- اثبات رابطه علی

اثبات تأثیر مستقیم پاسخ‌های هوش مصنوعی بر اقدام به خودکشی، به دلیل پیچیدگی عوامل روانی-اجتماعی، اغلب غیرممکن است (۲۶). در حقوق ایران، اصل ۱۵ قانون مسئولیت مدنی، تأکید دارد هر فردی که عملی خلاف قانون انجام دهد و خسارتی به دیگران وارد کند، موظف به جبران آن خسارت است (۲۷). این مورد را می‌توان به خطای الگوریتمی در سیستم‌های هوش مصنوعی نیز تعمیم داد.

ج-۳- شفافیت در مسئولیت حقوقی

در سیستم‌های مشاوره سنتی، درمانگران انسانی مسئولیت

اتخاذ کنند که منجر به تشخیص نادرست و مداخلات نامناسب شود (۱۹). این موضوع نشان می‌دهد که عدم توجه به تنوع فرهنگی و اجتماعی در طراحی سیستم‌های هوش مصنوعی می‌تواند پیامدهای جدی برای کاربران مختلف داشته باشد (۲۰).

ج-۶- ارزیابی دقت و صحت اطلاعات

ارزیابی دقت و اعتبار اطلاعات تولیدشده توسط سیستم‌های هوش مصنوعی، نیازمند داشتن تخصص در حوزه روانشناسی و مشاوره است. متخصصان پزشکی قانونی باید بتوانند با تحلیل عمیق و دقیق پاسخ‌های این سیستم‌ها، میزان اعتبار آنها را سنجیده و در خصوص صحت یا عدم صحت محتوای ارائه‌شده به کاربران (از جمله اطلاعات نادرست، گمراه‌کننده یا تهدیدآمیز) تشخیص‌های لازم را صادر کنند (۱۹). لذا تدوین چارچوب‌های حقوقی-پزشکی قانونی برای نظارت بر این فناوری‌ها، ضرورت تضمین دقت، ایمنی و پاسخگویی این فناوری در تعامل با کاربران را برجسته می‌کند.

ج-۷- چارچوب قانونی و نظارت جامع

عدم وجود چارچوب قانونی و نظارتی جامع در برنامه‌های روانشناسی و مشاوره مبتنی بر هوش مصنوعی، ابهامات زیادی را در مورد حقوق و مسئولیت‌های افراد درگیر در این سیستم‌ها ایجاد می‌کند (۱۶) در ایران نیز، عدم وجود چارچوب‌های قانونی و نظارتی برای استفاده از هوش مصنوعی در پزشکی و سلامت روان به یک چالش تبدیل شده است این خلأ قانونی می‌تواند به ارائه خدمات غیراستاندارد بینجامد و ضرورت همکاری وزارت بهداشت با مرکز ملی فضای مجازی برای تدوین پروتکل‌های نظارتی را تأکید می‌کند (۲۹). لذا تدوین چارچوب و قوانین با در نظر گرفتن سیستم نظارتی مستمر یکی از الزامات این حوزه می‌باشد.

یافته‌ها

مقاله حاضر با بررسی ابعاد مختلف استفاده از برنامه‌های روانشناسی و مشاوره مبتنی بر هوش مصنوعی در مشاوره خودکشی، نشان داد که اگرچه این فناوری پتانسیل‌های بسیاری برای بهبود دسترسی و کارایی خدمات سلامت روان دارد (۱،۲)، اما چالش‌های اخلاقی، حقوقی و حرفه‌ای را نیز به همراه دارد. این چالش‌ها عمدتاً ناشی از احتمال نقض اصول اساسی اخلاق حرفه‌ای (۸،۲۳) مسائل قانونی مربوط به مسئولیت پاسخگویی در قبال آسیب‌های ناشی از توصیه‌های نادرست این برنامه‌ها (۲۶) و محدودیت‌های این فناوری در درک پیچیدگی‌های عاطفی و فرهنگی و ارائه پاسخ‌های همدلانه هستند (۲۱).

یافته‌ها حاکی از این بود که لازم است در استفاده از فناوری‌های نوین در این حوزه، رویکردی محتاطانه و مسئولانه اتخاذ شود. به‌منظور بهره‌مندی از ظرفیت‌های هوش مصنوعی در این حوزه و کاهش خطرات احتمالی، می‌توان از این فناوری به عنوان ابزاری برای غربالگری اولیه، دریافت اطلاعات پایه و ارجاع به متخصصان انسانی استفاده شود، اما جایگزینی کامل آن با تخصص انسانی امکان‌پذیر نیست (۵) این رویکرد نه تنها با اصول اخلاقی منطبق است، بلکه به متخصصان انسانی اجازه می‌دهد تا با استفاده از دانش و تجربه خود، بهترین تصمیم‌ها را برای کمک به افراد در معرض خطر اتخاذ کنند (۱۴،۲۳) همچنین، تدوین استانداردهای اخلاقی و قانونی مشخص، توسعه سیستم‌های بومی و متناسب با فرهنگ و نظارت انسانی مستمر بر عملکرد سیستم‌های هوش مصنوعی ضروری است (۷) برای تقویت این چارچوب، لازم است سازوکارهای نظارتی مبتنی بر مشارکت نهادهای دانشگاهی و صنعتی توسعه یابد تا عملکرد الگوریتم‌ها به صورت مستمر ارزیابی شود (۲۲) همچنین، آموزش متخصصان سلامت روان در زمینه فناوری‌های نوین و تدوین پروتکل‌های استاندارد برای استفاده ایمن از هوش مصنوعی در بحران‌های روانی ضروری است (۱۶)

بهره‌برداری ایمن و اثربخش از هوش مصنوعی در پیشگیری از خودکشی نیازمند همکاری نزدیک متخصصان در حوزه‌های مختلف می‌باشد (۱۸). تشکیل کارگروه‌های تخصصی متشکل از روانپزشکان روانشناسان، مشاوران، حقوقدانان، اخلاق‌پژوهان و مهندسان فناوری اطلاعات، گامی اساسی در جهت کاهش ریسک‌های مرتبط با این فناوری و تضمین استفاده مسئولانه از آن است. در نهایت می‌توان گفت که هوش مصنوعی به تنهایی نمی‌تواند جایگزین همدلی انسانی، قضاوت بالینی و درک عمیق شرایط فردی شود (۲۱) لذا ضروری است که از این فناوری به عنوان ابزاری مکمل در چارچوبی اخلاقی و قانونی استفاده شود و همواره بر نقش محوری متخصصان انسانی تأکید شود (۴).

علاوه بر این، ایجاد پایگاه‌های داده ملی با رعایت اصول اخلاقی و حریم خصوصی، می‌تواند به بهبود دقت سیستم‌های هوش مصنوعی در شناسایی خطر خودکشی کمک کند (۲۵) در این راستا، همکاری بین‌المللی با سازمان‌هایی مانند سازمان جهانی بهداشت نیز برای تدوین استانداردهای جهانی و تطبیق آن‌ها با نیازهای محلی پیشنهاد می‌شود (۱۶).

بهره‌برداری ایمن و اثربخش از هوش مصنوعی در پیشگیری از خودکشی نیازمند همکاری نزدیک بین روانپزشکان، متخصصان اخلاق، مهندسان داده، و سیاستگذاران است (۱۳، ۲۲). این همکاری می‌تواند به توسعه چارچوب‌هایی بینجامد که داده‌های بالینی را با متغیرهای اجتماعی-فرهنگی تلفیق کرده و راه‌حل‌های نوینی برای چالش‌های موجود ارائه دهند (۲۰، ۲۵).

جدول ۱. چالش‌ها و راهکارهای پیشنهادی استفاده از برنامه‌های روانشناسی و مشاوره مبتنی بر هوش مصنوعی در مشاوره خودکشی

چالش‌ها	راهکارهای پیشنهادی
نقص در صلاحیت حرفه‌ای	طراحی چارچوب‌های نظارتی برای مجوزدهی به سیستم‌های هوش مصنوعی با همکاری سازمان‌های ذیربط از جمله سازمان نظام روانشناسی و مشاوره، وزارت بهداشت و با نظارت مستمر متخصصان
خطر آسیب‌رسانی	نظارت مستمر متخصصان بر عملکرد سیستم‌ها
شائبه در رضایت آگاهانه	تدوین قوانین الزام‌آور جهت ارائه توضیحات شفاف و دریافت رضایت آگاهانه از کاربران پیش از استفاده از سیستم‌ها، مطابق با استانداردهای ملی و بین‌المللی
ابهام در مسئولیت‌پذیری	تدوین استانداردهای اخلاقی و حقوقی برای تعیین مسئولیت مشترک میان توسعه‌دهنده، ارائه‌دهنده و کاربر، مبتنی بر سهم و نقش هر یک در فرآیند طراحی و استفاده از سیستم‌ها
نقض حریم خصوصی	تطابق با قوانین ملی (مانند ماده ۷۵ قانون جرائم رایانه‌ای ایران) و ذخیره داده‌ها در سرورهای داخلی
سوگیری داده‌های آموزشی	تولید پایگاه داده‌های بومی، آموزش تدوین الگوریتم‌ها با حساسیت‌های فرهنگی بر اساس استانداردهای بین‌المللی
ناسازگاری فرهنگی	توسعه الگوریتم‌های مبتنی بر داده‌های بومی و حساس به فرهنگ و ایجاد سازوکارهای نظارتی مبتنی بر مشارکت نهادهای دانشگاهی و صنعتی
عدم وجود چارچوب قانونی و نظارتی	همکاری وزارت بهداشت با مرکز ملی فضای مجازی برای تدوین قوانین شفاف و پروتکل‌های نظارتی جهت استفاده از برنامه‌های روانشناسی و مشاوره مبتنی بر هوش مصنوعی

جدول ۲. محدودیت‌های ذاتی و راهکارهای پیشنهادی استفاده از برنامه‌های روانشناسی و مشاوره مبتنی بر هوش مصنوعی در مشاوره خودکشی

راهکارهای پیشنهادی	محدودیت‌های ذاتی
مانند تغییرات رفتاری	تلفیق فناوری با ارزیابی بالینی متخصصان حوزه سلامت جهت تحلیل نشانه‌های غیرکلامی مانند لحن صدا، زبان بدن و علائم روان‌شناختی
استفاده از هوش مصنوعی به عنوان ابزار مکمل برای غربالگری اولیه و ارجاع به متخصصان حوزه سلامت در شرایط بحرانی	امکان شکست در مدیریت بحران
ترکیب هوش مصنوعی با مداخله متخصصان جهت درک پیچیدگی‌های عاطفی، فرهنگی و اجتماعی	وابستگی به داده‌های کمی

نتیجه‌گیری

فرهنگ، ایجاد چارچوب‌های نظارتی برای کاهش سوگیری‌های سیستماتیک، و تقویت همکاری بین‌رشته‌ای میان حوزه‌های روانپزشکی، پزشکی قانونی، حقوق و فناوری، از الزامات اساسی برای تضمین اثربخشی و عدالت در کاربرد این سیستم‌هاست. در نهایت، موفقیت این فناوری در گروهی حفظ تعادل میان نوآوری و اصول اخلاقی، با تأکید بر نقش محوری متخصصان در تصمیم‌گیری‌های حیاتی می‌باشد.

محدودیت‌های مطالعه

از محدودیت‌های این پژوهش می‌توان به تعمیم‌پذیری محدود نتایج به دلیل تک‌موردی بودن مطالعه، عدم دسترسی به داده‌های آموزشی الگوریتم‌های تجاری، و چالش‌های فرهنگی-اجتماعی خاص ایران اشاره کرد که ممکن است در سایر بسترها مصداق نداشته باشد.

برنامه‌های روانشناسی و مشاوره مبتنی بر هوش مصنوعی در مشاوره خودکشی، با وجود مزایای بالقوه در گسترش دسترسی به خدمات سلامت روان، با چالش‌های چندوجهی در سطوح مختلف مواجه است. این فناوری، علیرغم پیشرفت‌های فنی، فاقد قابلیت درک پیچیدگی‌های عاطفی و فرهنگی است که برای ارزیابی دقیق خطر خودکشی حیاتی محسوب می‌شوند. علاوه بر این، ابهام در تعیین مسئولیت حقوقی ناشی از خطاهای الگوریتمی و نگرانی‌های مربوط به حریم خصوصی داده‌ها، استفاده از آن را در موقعیت‌های بحرانی با محدودیت‌های جدی روبه‌رو می‌سازد. راهکارهای کلیدی برای بهره‌برداری ایمن و مسئولانه از این فناوری، شامل تعریف نقش مکمل هوش مصنوعی در غربالگری اولیه و تسهیل ارجاع به متخصصان انسانی، همراه با تدوین استانداردهای اخلاقی-حقوقی شفاف است. توسعه الگوریتم‌های مبتنی بر داده‌های بومی و حساس به

References

1. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J*. 2019;6(2):94-98.
2. Graham S, Depp C, Lee EE, Nebeker C, Tu X, Kim HC, et al. Artificial intelligence for mental health and mental illnesses: an overview. *Curr Psychiatry Rep*. 2019;21(11):116.
3. D'Alfonso S. AI in mental health. *Curr Opin Psychol*. 2020;36(0):112-117.
4. Fiske A, Henningsen P, Buyx A. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *J Med Internet Res*. 2019;21(5):e13216.
5. Belsher BE, Smolenski DJ, Pruitt LD, Bush NE, Beech EH, Workman DE, et al. Prediction models for suicide attempts and deaths: a systematic review and simulation. *JAMA Psychiatry*. 2019;76(6):642-651.
6. Char DS, Shah NH, Magnus D. Implementing machine learning in health care—addressing ethical challenges. *N Engl J Med*. 2018;378(11):981-983.
7. Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L. The ethics of algorithms: mapping the debate. *Big Data Soc*. 2016;3(2):1-15.
8. Martinez-Martin N, Kreitmair K. Ethical issues for direct-to-consumer digital psychotherapy apps: addressing accountability, data protection, and consent. *JMIR Ment Health*. 2018;5(2):e32.
9. Iranian Psychological and Counseling Association System of the Islamic Republic of Iran. Code of ethics [Internet]. 2017 [cited 2023 Apr 08]. [Available from: <https://pcoiran.ir/media/2023/1/10/1673338743108280.pdf>]. [Persian]
10. American Psychological Association. Ethical principles of psychologists and code of conduct [Internet]. Washington, DC: APA; 2017 [cited 2023 Oct 5]. [Available from: <https://www.apa.org/ethics/code/ethics-code-2017.pdf>].
11. British Psychological Society. Code of ethics and conduct [Internet]. Leicester: BPS; 2021 [cited 2025 Apr 8]. Available from: <https://www.bps.org.uk/news-and-policy/bps-code-ethics-and-conduct>.
12. Canadian Psychological Association. Canadian code of ethics for psychologists [Internet]. 4th ed. Ottawa, ON: CPA; 2017 [cited 2025 Apr 8]. [Available from: https://cpa.ca/docs/File/Ethics/CPA_Code_2017_4thEd.pdf].
13. Gerke S, Minssen T, Cohen G. Ethical and legal challenges of artificial intelligence-driven healthcare. In: Bohr A, Memarzadeh K, editors. *Artificial intelligence in healthcare*. Cambridge: Academic Press; 2020. p. 295-336.
14. Islamic Republic of Iran. Islamic Penal Code [Qānūn-e Mojāzāt-e Eslāmī], Articles 495-497. 2013 May 22 [cited 2023 Oct 10]. Available from: <https://rc.majlis.ir/fa/law/show/845048>. [Persian]
15. Franklin JC, Ribeiro JD, Fox KR, Bentley KH, Kleiman EM, Huang X, et al. Risk factors for suicidal thoughts and behaviors: a meta-analysis of 50 years of research. *Psychol Bull*. 2017 Feb;143(2):187-232.
16. Wells C, Hill S, Argáez C. Artificial Intelligence and Machine Learning in Mental Health Services: A Literature Review. Ottawa: Canadian Agency for Drugs and Technology in Health (CADTH), Mental Health Commission of Canada (MHCC); 2021 June. [Available from: https://www.cda-amc.ca/sites/default/files/attachments/2021-06/artificial_intelligence_and_machine_learning_in_mental_health_services_literature_review.pdf]
17. Steeg S, Quinlivan L, Nowland R, Carroll R, Casey D, Clements C, Cooper J, Davies L, Knipe D, Ness J, O'connor RC. Accuracy of risk scales for predicting repeat self-harm and suicide: a multicentre, population-level cohort study using routine clinical data. *BMC psychiatry*. 2018 Dec;18:1-1.
18. Simon GE. Evidence-based management of depression. *N Engl J Med*. 2017 Jul;377(4):345-55.
19. O'Neil C. *Weapons of math destruction: how big data increases inequality and threatens democracy*. New York: Broadway Books; 2016.
20. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K. and Galstyan, A., 2021. A survey on bias and

- fairness in machine learning. ACM computing surveys (CSUR), 54(6), pp.1-35. [Persian]
21. Elyoseph Z, Levkovich I. Beyond human expertise: the promise and limitations of ChatGPT in suicide risk assessment. *Front Psychiatry*. 2023 May;14:1213141.
 22. Smith J, Brown T, Lee R. Hybrid AI-human systems in mental health. *J Med Ethics*. 2023 Mar;49(3):123-30.
 23. Beauchamp TL, Childress JF. Principles of biomedical ethics. 8th ed. New York: Oxford University Press; 2019.
 24. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. *Nat Mach Intell*. 2019;1(9):389-399.
 25. Abbas M, Timuri M. A review of ethical and legal challenges of artificial intelligence applications in the healthcare system. *J Med Ethics Res*. 2024;17(48):1-11. [Available from: <https://doi.org/10.22037/mej.v17i48.44053>]. [Persian]
 26. Price WN, Cohen IG, Lehmann H. Big data, health law, and bioethics. Cambridge: Cambridge University Press; 2016.
 27. Islamic Republic of Iran. Civil Liability Act [Internet]. Article 15. Enacted 1960 Mar 28 [cited 2025 Apr 16]. [Available from: <https://rc.majlis.ir/fa/law/>]. [Persian]
 28. Islamic Republic of Iran. Computer Crimes Law Article 75 [Internet]. 2009 [cited 2025 Apr 8]. [Available from: <https://rc.majlis.ir/fa/law/show/135717>]. [Persian]
 29. Abbas M, Timuri M. A review of ethical and legal challenges of artificial intelligence applications in the healthcare system. *J Med Ethics Res*. 2024;17(48):1-11. [Available from: <https://doi.org/10.22037/mej.v17i48.44053>]. [Persian]

Examining the Challenges of Employing Artificial Intelligence in Suicide Counseling Through an Interdisciplinary Approach

Manzara Amirkhani^{1*}, Mohammad Hosein Sadeghian², Ahmad Hajebi³, AmirAhmad Shojaei^{4,5}

Abstract

Objective: The rapid proliferation of artificial intelligence (AI) in mental health services, particularly in sensitive domains such as suicide counseling, has raised significant ethical, clinical, and legal challenges. This study aimed to analyze user interactions with AI-based psychological and counseling applications, identify key challenges in ethical, psychiatric, legal, and forensic medical domains, and propose strategies for the responsible utilization of this technology in suicide prevention.

Methods and Materials: This qualitative study employed a real-world clinical scenario (an interaction between an individual with suicidal ideation and an AI-based counseling application), analyzed Iranian legal documents (the Islamic Penal Code and the Cybercrimes Law), psychological ethics standards (Iranian, European, and American), and domestic and international research. Data included recorded user conversations, algorithmic responses, and behavioral outcome assessments. Data analysis was conducted through an interdisciplinary approach (Ethical, Psychiatry, Forensic Medicine and Legal) using qualitative content analysis and comparative methods.

Results: AI-based counseling systems lack the professional competence to replace clinical judgment by human experts and demonstrate an inability to comprehend emotional, cultural, and social nuances. Non-empathetic algorithmic responses may exacerbate suicide risk. Key challenges include privacy violations, cultural incompatibility, ambiguity in legal accountability, and algorithmic bias.

Conclusion: The use of AI in suicide counseling necessitates a cautious, responsible approach grounded in human expert oversight. Such systems may serve as complementary tools for preliminary screening and urgent referrals to human professionals. The development of comprehensive ethical-legal frameworks, culturally sensitive and context-aware algorithms, interdisciplinary collaboration (psychiatry, law, and technology), and robust regulatory mechanisms to ensure data privacy and security are imperative. Additionally, establishing clear performance standards and rigorous monitoring protocols for these systems is essential.

Keywords: Suicide, Counseling, Ethical Challenges, Artificial Intelligence, Interdisciplinary Collaboration

1*. Corresponding Author: PhD in Counseling, Deputy for Education, Ministry of Health and Medical Education, Tehran, Iran
dr.mnzs.amirkhani@gmail.com

2. Associate Professor, Department of Forensic Medicine, School of Medicine Tehran University of Medical Sciences Tehran, Iran

3. Professor, Department of Psychiatry, Research Center for Addiction & Risky Behaviors (ReCARB), Iran University of Medical Sciences, Tehran, Iran

4. Associate Professor, Medical Ethics and History of Medicine Research Center, Tehran University of Medical Sciences, Tehran, Iran

5. Associate Professor, Department of Medical Ethics, School of Medicine, Tehran University of Medical Sciences Tehran, Iran