تأثیرات هوش مصنوعی در ارزیابی عملکرد کارکنان: فرصت‌ها، چالش‌ها و الزامات حکمرانی الگوریتمی

فاطمه عطری[۱] و منصوره کریمیان[۲]

تاریخ دریافت: ۱۴۰۴/۰۳/۲۹ پذیرش نهایی: ۱۴۰۴/۱۰/۰۱

چکیده

گسترش هوش مصنوعی در نظام‌های مدیریت عملکرد، الگوهای سنتی ارزیابی کارکنان را دگرگون کرده و آن را به فرآیندی داده‌محور، پیوسته و پیش‌بینانه تبدیل کرده است. با وجود مزایای قابل توجه این فناوری، از جمله افزایش دقت تحلیل‌های منابع انسانی و کاهش سوگیری‌های انسانی، پیامدهای رفتاری و ادراکی آن هنوز به‌طور جامع تبیین نشده است. پژوهش حاضر با هدف تحلیل تأثیر مؤلفه‌های دقت ادراک‌شده، شفافیت الگوریتمی و عدالت ادراک‌شده بر اعتماد و پذیرش سیستم‌های ارزیابی عملکرد مبتنی بر هوش مصنوعی انجام شده است. این مطالعه از روش‌شناسی ترکیبی بهره برده است؛ در بخش کیفی، داده‌ها از طریق تحلیل مضمون مصاحبه با خبرگان استخراج شد و در بخش کمی، مدل مفهومی با استفاده از داده‌های شبیه‌سازی‌شده ($N=300$) و روش مدل‌یابی مسیر، تست گردید. یافته‌ها نشان می‌دهد که هر سه سازه ادراکی، اثر مثبت و معناداری بر اعتماد دارند که در این میان شفافیت الگوریتمی قوی‌ترین پیش‌بین است($\beta=0.35, p<0.001$) . علاوه بر این، مدل توانست ۶۲ درصد از واریانس اعتماد را تبیین کند. همچنین اعتماد، اثر قابل‌توجهی بر پذیرش فناوری دارد ($\beta=0.81, p<0.001$) و ۶۵ درصد از واریانس آن را تبیین می‌کند. نتایج بوت‌استرپ نیز اثرات غیرمستقیم مثبت و معنادار دقت، شفافیت و عدالت را بر پذیرش فناوری از طریق اعتماد تأیید کرد. این یافته‌ها بیانگر آن است که موفقیت هوش مصنوعی در ارزیابی عملکرد نه‌تنها به قابلیت‌های فنی، بلکه به کیفیت حکمرانی الگوریتمی و ادراک کارکنان از شفافیت، انصاف و قابل‌اتکا بودن سیستم بستگی دارد. پژوهش حاضر با ارائه شواهد تجربی، اهمیت طراحی مدل‌های هیبریدی انسان–هوش مصنوعی و توسعه چارچوب‌های حکمرانی الگوریتمی را برجسته می‌کند و مسیر پژوهش‌های آینده را به سمت تلفیق عوامل فنی، اخلاقی و روان‌شناختی هدایت می‌نماید.

کلمات کلیدی: : هوش مصنوعی، ارزیابی عملکرد کارکنان، شفافیت الگوریتمی، عدالت ادراک‌شده، اعتماد به سیستم‌های هوشمند

[۱] دانشگاه آزاد اسلامی واحد تهران جنوب، دانشکده علوم تربیتی، ایران، تهران
[۲] دانشگاه آزاد اسلامی واحد تهران جنوب، دانشکده علوم تربیتی، ایران، تهران

مقدمه

تأثیرات هوش مصنوعی در ارزیابی عملکرد کارکنان تحول دیجیتال در دهه اخیر، ماهیت مدیریت منابع انسانی را به‌گونه‌ای دگرگون ساخته که فرایندهای سنتی مبتنی بر قضاوت انسانی، جای خود را به تصمیم‌سازی‌های داده‌محور و الگوریتمی داده‌اند(Barocas & Selbst, 2016). در این میان، ارزیابی عملکرد کارکنان به‌عنوان یکی از حساس‌ترین و راهبردی‌ترین فرایندهای سازمانی، بیشترین تأثیر را از گسترش فناوری‌های نوظهور به‌ویژه هوش مصنوعی پذیرفته است(Chen, Chen, & Lin, 2020). سازمان‌های پیشرو امروزی به جای اتکا بر برداشت‌های ذهنی مدیران و جلسات دوره‌ای[1]، از تحلیل کلان‌داده‌های رفتاری، مدل‌های یادگیری ماشین، پردازش زبان طبیعی و سیستم‌های پیش‌بینانه بهره می‌گیرند تا کیفیت سنجش عملکرد نیروی انسانی را افزایش دهند و تصمیمات مبتنی بر شواهد را جایگزین قضاوت‌های فردی کنند(Malik, 2024).

تغییر الگوی ارزیابی عملکرد از رویکرد پس‌نگر و توصیفی به الگوی پیش‌نگر، بی‌درنگ و تحلیلی، طیف گسترده‌ای از دستاوردها را برای سازمان‌ها به همراه داشته است(Basnet, 2024). افزایش دقت اندازه‌گیری عملکرد، شناسایی کارکنان با پتانسیل بالا، ارائه بازخوردهای شخصی‌سازی‌شده، بهبود برنامه‌ریزی جانشین‌پروری و پیش‌بینی ریسک‌های منابع انسانی از جمله مهم‌ترین مزایای گزارش‌شده در پژوهش‌های بین‌المللی هستند(Căvescu & Popescu, 2025). در واقع، استفاده از هوش مصنوعی سبب می‌شود عملکرد شغلی نه‌تنها به‌عنوان خروجی‌های قابل مشاهده، بلکه به‌عنوان شبکه‌ای از رفتارهای فردی و تعاملی که در پلتفرم‌ها و سیستم‌های سازمانی ردپای دیجیتال می‌گذارند، قابل سنجش باشد(Formosa, Bankins, Griep, & Richards, 2025; Shin, Choi, & Kim, 2025). این تحول، امکان تحلیل‌های عمیق‌تر از جمله پیش‌بینی شاخص‌هایی مانند احتمال ترک خدمت، کاهش درگیری شغلی یا افت کیفی عملکرد را فراهم می‌سازد و سازمان را در اتخاذ تصمیمات پیشگیرانه توانمند می‌سازد(Boroumandzadeh, Pirzad, Nikbakhsh, & Soltani, 2025).

با وجود این فرصت‌ها، نگرانی‌های گسترده‌ای نیز درباره پیامدهای اخلاقی و اجتماعی این فناوری مطرح شده است. استفاده از داده‌های کلان رفتاری، تحلیل مکالمات و نظارت مستمر دیجیتال، مرزهای حریم خصوصی کارکنان را مبهم می‌سازد و می‌تواند باعث شکل‌گیری احساس «نظارت‌زدگی» و تهدید کرامت انسانی شود. علاوه‌براین، الگوریتم‌ها در معرض تبعیض داده‌محور قرار دارند؛ به این معنا که اگر داده‌هایی که مدل بر مبنای آن‌ها آموزش دیده، حاوی سوگیری‌های تاریخی باشد، این سوگیری‌ها در مقیاسی بزرگ بازتولید خواهند شد. مطالعات نشان داده‌اند که برخی مدل‌های ارزیابی مبتنی بر زبان طبیعی نسبت به جنسیت، سبک‌های ارتباطی یا فرهنگ‌های خاص حساسیت نامتوازن دارند و در نتیجه می‌توانند موجب نابرابری در امتیازدهی عملکرد شده و عدالت سازمانی را خدشه‌دار کنند(Khanal, 2026; Mask & Pearl, 2024; Yaqoob & Robbins, 2024).

---

[1] Performance Review

چالش مهم دیگر، ابهام‌پذیری الگوریتمی است. تصمیمات مبتنی بر هوش مصنوعی، به‌ویژه در مدل‌های پیچیده یادگیری عمیق، همواره قابلیت توضیح کافی برای کارکنان و مدیران ندارند. در چنین شرایطی، کارکنان ممکن است درک نکنند چگونه، چرا و بر اساس کدام شواهد عملکرد آنان ارزیابی شده است. این وضع، اعتماد سازمانی را تضعیف کرده و پذیرش فناوری را کاهش می‌دهد.

از آنجا که ارزیابی عملکرد مستقیماً با ارتقاء، پرداخت‌ها و آینده شغلی افراد در ارتباط است، عدم شفافیت در این حوزه می‌تواند تبعات حقوقی، روانی و فرهنگی قابل توجهی ایجاد کند(Al-Sulaiti et al., 2024; Liem et al., 2018; Mittelstadt, Russell, & Wachter, 2019).

به همین دلیل، امروزه در سطح جهانی بحث‌هایی مانند «حکمرانی الگوریتمی»، «شرح‌پذیری هوش مصنوعی» و «عدالت الگوریتمی» به یکی از محورهای اصلی مطالعات مدیریت و حقوق فناوری تبدیل شده است. بسیاری از سازمان‌ها و کشورها در حال تدوین چارچوب‌های اخلاقی و قانونی برای پیاده‌سازی مسئولانه هوش مصنوعی در محیط کار هستند؛ اما این چارچوب‌ها هنوز در مرحله شکل‌گیری بوده و استانداردهای جامعی که بتواند نیازهای متنوع سازمان‌ها را پوشش دهد، به‌طور کامل توسعه نیافته است.

از سوی دیگر، بخش قابل توجهی از پژوهش‌های موجود به مزیت‌های فناورانه یا ابعاد فنی مدل‌ها تمرکز داشته و کمتر به تجربه کارکنان، ادراک عدالت، اعتماد سازمانی و تبعات رفتاری توجه کرده‌اند. به‌ویژه در زمینه ترکیب داده‌های رفتاری واقعی با شاخص‌های روان‌شناختی (مانند درگیری شغلی، احساس انصاف و اعتماد به سیستم)، خلأ پژوهشی معناداری مشاهده می‌شود؛ بنابراین، نیاز به یک تحلیل جامع وجود دارد که ضمن بررسی دستاوردهای فناورانه، تأثیرات انسانی و سازمانی این تحول را نیز به‌طور هم‌زمان واکاوی نماید.

در این چارچوب، پژوهش حاضر با تمرکز بر نقاط تماس بین فناوری، رفتار انسانی و سیاست‌های منابع انسانی تلاش می‌کند تصویری شفاف از مزایا، مخاطرات و الزامات کاربرد هوش مصنوعی در ارزیابی عملکرد کارکنان ارائه دهد. هدف نهایی پژوهش، ارائه یک چارچوب حکمرانی الگوریتمی است که بتواند توسعه‌پذیری و مشروعیت اجتماعی این فناوری را تضمین کند؛ به‌گونه‌ای که دقت ارزیابی و عدالت ادراک‌شده به‌عنوان دو ستون اصلی اعتماد سازمانی تقویت شوند و سازمان قادر باشد ضمن بهره‌گیری از هوش مصنوعی، اصول اخلاق حرفه‌ای و حقوق فردی کارکنان را حفظ نماید.

بنابراین، پرسش اصلی این است که چگونه می‌توان از ظرفیت‌های هوش مصنوعی برای بهبود ارزیابی عملکرد استفاده کرد، بدون آنکه عدالت، شفافیت و امنیت روانی کارکنان آسیب ببیند؟ پاسخ به این پرسش نه‌تنها برای مدیران و سیاست‌گذاران سازمانی، بلکه برای جامعه دانشگاهی و کارشناسان فناوری اطلاعات نیز اهمیت حیاتی دارد. این پژوهش با اتخاذ رویکردی تحلیلی و میان‌رشته‌ای، تلاش دارد راهکارهایی کاربردی و مبتنی بر شواهد برای پیاده‌سازی مسئولانه ارزیابی عملکرد مبتنی بر هوش مصنوعی ارائه نماید و مسیر تحقیقات آینده در این حوزه را روشن‌تر سازد.

۲. ادبیات و پیشینه تحقیق

تحول دیجیتال، به‌ویژه پس از ورود فناوری‌های هوشمند به محیط‌های سازمانی، موجب شده فرایند ارزیابی عملکرد که زمانی ماهیتی ایستا و پس‌نگر داشت، اکنون به رویکردی پویا، پیش‌بینانه و داده‌محور تبدیل شود. در این الگو، ارزیابی عملکرد نه‌تنها مبتنی بر نتایج قابل مشاهده، بلکه متکی بر ردپای دیجیتال رفتارهای شغلی است که در سامانه‌های سازمانی به‌جا می‌ماند(Tambe, Cappelli, & Yakubovich, 2019). این دگرگونی، با ظهور مفهوم «تحلیل سرمایه انسانی» همراه است؛ مفهومی که هدف آن تبدیل داده‌های رفتاری به بینش‌های مدیریتی قابل استفاده است. بنابراین، مدیریت عملکرد در عصر جدید بیش از آنکه درباره قضاوت فردی مدیران باشد، درباره شناخت علمی توانمندی‌ها و الگوهای رشد کارکنان است. این وضعیت، نیاز سازمان‌ها به الگوریتم‌هایی را افزایش داده که قادر به تحلیل حجم انبوهی از داده‌ها و ارائه تصمیماتی دقیق و به‌موقع باشند(Bashynska, Prokopenko, & Sala, 2023).

هوش مصنوعی در سه محور اصلی فرایند ارزیابی عملکرد را غنی‌سازی کرده است که عبارتند از: تحلیل داده‌های متنوع و چندمنبعی

مانند پیام‌ها، زمان‌بندی کارها، داده‌های سنسورها و گزارش‌های سیستمی. پژوهش‌ها نشان می‌دهند که آنالیز HR مبتنی بر AI می‌تواند دقت در ارزیابی عملکرد را تا ۳۰ درصد افزایش دهد و فرایندهای تصمیم‌گیری منابع انسانی مثل ارتقاء و جانشین‌پروری را بهبود بخشد (Mitchell & Brynjolfsson, 2023). بنابراین، ارزیابی از یک فعالیت اداری به ابزاری استراتژیک برای خلق مزیت رقابتی تبدیل شده است.

درحالی‌که مزایای فناورانه قابل توجه است، پژوهش‌ها به‌طور جدی نسبت به پیامدهای انسانی هشدار می‌دهند. یکی از مهم‌ترین ریسک‌ها تبعیض الگوریتمی است. الگوریتم‌هایی که بر داده‌های تاریخی آموزش می‌بینند، ممکن است سوگیری‌های نهادینه‌شده را بازتولید کنند (Barocas & Selbst, 2016). برای مثال، اگر سازمانی در گذشته تمایل به ترفیع مردان داشته باشد، مدل پیش‌بین نیز ممکن است جنسیت را به‌عنوان متغیری اثرگذار بر عملکرد فرض کند. چالش دیگر تشدید نظارت سازمانی است. کارمند در محیطی قرار می‌گیرد که هر اقدام، گفتار و تصمیم او ممکن است تحلیل شود و این احساس کنترل دائمی، امنیت روانی و اعتماد سازمانی را تضعیف می‌کند. همچنین ابهام‌پذیری یا عدم شرح‌پذیری بسیاری از الگوریتم‌ها، پذیرش فناوری را تهدید می‌کند. کارکنان انتظار دارند منطق ارزیابی عملکردشان قابل پرسش و بررسی باشد؛ عدم شفافیت می‌تواند به باورهای منفی نسبت به عدالت ارزیابی و در نهایت کاهش تعهد سازمانی منجر شود (Wang & Siau, 2019).

گسترش فناوری‌های هوش مصنوعی موجب تغییرات بنیادین در فرایند ارزیابی عملکرد شده است و پژوهش‌های متعددی در سطح بین‌المللی به بررسی ابعاد مختلف آن پرداخته‌اند. برای نمونه، (Tambe et al. 2019) در مطالعه‌ای جامع دریافتند که سیستم‌های هوش مصنوعی با تحلیل حجم بالای داده‌های رفتاری کارکنان، دقت ارزیابی را افزایش داده و سوگیری‌های انسانی را کاهش می‌دهند. در ادامه این خط پژوهشی، وانگ و همکاران (Wang & Siau, 2019) بیان کردند که آنچه موجب پذیرش سیستم‌های ارزیابی هوشمند می‌شود نه فقط دقت الگوریتمی، بلکه دقت ادراک‌شده توسط کارکنان است؛ یعنی برداشت ذهنی افراد از میزان صحت تصمیمات سیستم، نقش تعیین‌کننده‌ای در اعتماد آنان دارد.

از منظر دیگر، هانگ و همکاران (Huang and Rust 2022) تأکید کردند که هوش مصنوعی با ارتقای تحلیل‌های داده‌محور، می‌تواند مدیریت عملکرد را از رویکردی واکنشی به رویکردی پیش‌بینانه تبدیل کند. این موضوع باعث بهبود تصمیم‌گیری مدیران و ارتقای عملکرد سازمانی می‌شود. همچنین، مایکل و همکارانش (Mitchell & Brynjolfsson, 2023) نشان دادند سازمان‌هایی که از سیستم‌های هوشمند ارزیابی استفاده می‌کنند، بهبود قابل توجهی در عملکرد کارکنان و تحقق اهداف مشاهده می‌کنند، مشروط بر اینکه کارکنان فرایند تصمیم‌گیری الگوریتم را قابل‌اعتماد بدانند.

یکی از چالش‌های قابل‌توجه در این حوزه، مسئله شفافیت و شرح‌پذیری تصمیمات الگوریتمی است. ژانگ و همکاران در پژوهشی تجربی ثابت کردند که زمانی کارکنان منطق امتیازدهی سیستم را درک کنند، احساس اطمینان و پذیرش فناوری افزایش می‌یابد(Zhang et al., 2021). یافته‌های سامک و مولر نیز حاکی از آن است که نبود شفافیت می‌تواند زمینه‌ساز مقاومت کارکنان و کاهش انگیزه شود، حتی اگر سیستم از نظر آماری دقیق باشد(Samek & Müller, 2020) .

موضوع عدالت ادراک‌شده نیز بخش مهمی از ادبیات را تشکیل می‌دهد. باروکاس و همکاران هشدار دادند که اگرچه هدف از به‌کارگیری الگوریتم‌ها کاهش تبعیض است، اما داده‌های ناقص یا سوگیری‌های پنهان می‌توانند به بی‌عدالتی الگوریتمی منجر شوند(Barocas & Selbst, 2016). این یافته با پژوهش Ajunwa(2017) سازگار است که نشان داد استفاده ناآگاهانه از AI در HR می‌تواند تبعیض‌های جدیدی ایجاد کند.

در خصوص «اعتماد به فناوری»، گیلسون و والی بیان کردند که اعتماد به سیستم هوشمند، میانجی کلیدی بین کیفیت ارزیابی و پذیرش کاربران است(Glikson and Woolley, 2020). در همین زمینه، مک‌برد دریافت که افزایش شفافیت و ارتباطات انسانی در کنار الگوریتم‌ها، اعتماد کارکنان و رفتارهای شغلی مثبت را تقویت می‌کند(McBrid, 2021).

بررسی ادبیات داخلی نشان می‌دهد که استفاده از فناوری‌های هوشمند در فرایندهای منابع انسانی در ایران نیز مورد توجه پژوهشگران قرار گرفته است؛ هرچند هنوز در مراحل اولیه بلوغ است.

رضایی و همکاران (۱۳۹۹) در یک مطالعه تجربی بیان کردند که دیجیتالی شدن ارزیابی عملکرد موجب کاهش خطاهای انسانی و افزایش شفافیت می‌شود و در نهایت سطح رضایت کارکنان را ارتقا می‌دهد. این یافته با پژوهش آشنا و امیری (۱۴۰۲) همراستا است که نشان داد استفاده از سیستم‌های دیجیتال عملکرد کارکنان را معنادار افزایش می‌دهد.

در حوزه عدالت ادراک‌شده، کریمی و داوری (۱۴۰۰) دریافتند که عدم توضیح‌پذیری تصمیمات سامانه‌های هوشمند می‌تواند ادراک کارکنان از عدالت را تضعیف کند. ایزدی و همکاران (۱۴۰۱) نیز تاکید کردند که ادراک کارکنان از انصاف سیستم، عامل حیاتی برای پذیرش و اعتماد آنان به فناوری است.

در زمینه دقت ادراک‌شده، محمدی و شریفی (۱۴۰۰) نشان دادند که کارکنان زمانی سیستم ارزیابی هوشمند را دقیق می‌دانند که داده‌های مورد استفاده در ارزیابی جامع و به‌روز باشد. علاوه‌براین، عظیمی و حسینی (۱۳۹۸) اشاره کردند که اگر کارکنان احساس کنند اطلاعات شخصی آن‌ها نادرست یا ناقص است، احتمال عدم پذیرش سیستم هوشمند افزایش می‌یابد.

پژوهش‌های دیگری نیز در کشور به موضوع اعتماد به فناوری پرداخته‌اند. برای نمونه، مرادی و صادقی (۱۳۹۹) بیان کردند که اعتماد به سیستم‌های هوشمند زمانی ایجاد می‌شود که کارکنان تجربه مثبت و دسترسی شفاف به نتایج ارزیابی داشته باشند. همچنین، رستمی و پاک‌سرشت (۱۴۰۰) نشان دادند که برگزاری دوره‌های آموزش آشنایی با فناوری و ایجاد امکان اعتراض نسبت به نتایج، منجر به افزایش اعتماد کارکنان می‌شود.

مرور این پژوهش‌ها نشان می‌دهد که ارتباط میان دقت، شفافیت، عدالت ادراک‌شده و اعتماد در نهایت بر عملکرد شغلی تأثیر می‌گذارد؛ با این حال، همگرایی این متغیرها در یک مدل جامع کمتر مورد توجه قرار گرفته است که همین مسئله شکاف نظری مهمی برای پژوهش حاضر ایجاد می‌کند.

بر اساس مرور ادبیات، سه نظریه اصلی برای تحلیل رویکرد AI در ارزیابی عملکرد مناسب است:

الف- نظریه عدالت سازمانی جهت توضیح اثر ارزیابی الگوریتمی بر ادراک انصاف

ب- نظریه شرح‌پذیری هوش مصنوعی جهت تبیین نیاز کاربران به منطق تصمیمات الگوریتمی

ج- نظریه پذیرش فناوری: بررسی عوامل مؤثر بر پذیرش یا مقاومت کارکنان

این نظریه‌ها چارچوبی فراهم می‌کنند تا پیامدهای فناورانه نه‌فقط از منظر فنی، بلکه از منظر انسان‌محور تحلیل شوند.

پیشینه تحقیق نشان می‌دهد که سوگیری در الگوریتم‌ها می‌تواند از کد، داده یا تفسیر انسانی ناشی شود. این سوگیری‌ها گاه منجر به تصمیمات خودکاری می‌شوند که پیامدهای واقعی بر زندگی حرفه‌ای افراد دارد؛ از عدم ارتقاء گرفته تا بی‌عدالتی در پاداش. بنابراین، عدالت سازمانی در عصر الگوریتم دیگر یک بحث اخلاقی صرف نیست؛ به موضوعی حقوقی و حاکمیتی تبدیل شده است.

مطالعات موجود اگرچه به ابعاد مهمی پرداخته‌اند، اما چند خلأ اساسی باقی مانده است. اولین خلأ آن کمبود پژوهش‌های ترکیبی که داده‌های رفتاری واقعی را با شاخص‌های روان‌شناختی (اعتماد، امنیت روانی، درگیری شغلی) پیوند دهد. مورد دیگر فقدان مدل‌های یکپارچه است که همزمان نمی‌توانند سه معیار دقت الگوریتمی، عدالت ادراک‌شده و حفاظت از حریم خصوصی را بررسی کنند. نهایتاً محدود بودن مطالعات در بستر فرهنگی ایران است که به تفاوت‌های بنیادین در قوانین کار، ارزش‌ها و سبک ارتباطی بستگی دارد. در نتیجه، محققان توصیه کرده‌اند که نسل جدید پژوهش‌ها باید بینش فنی و انسانی را ادغام کند تا راهکارهایی متوازن برای سازمان‌ها فراهم شود.

در نهایت ذکر این نکته ضروری است که مرور ادبیات تحقیق نشان می‌دهد هوش مصنوعی ارزیابی عملکرد را کارآمدتر، دقیق‌تر و آینده‌نگرتر کرده است، اما بدون حکمرانی الگوریتمی، می‌تواند به بی‌عدالتی و بی‌اعتمادی بینجامد. بنابراین نیاز است نگاه از «اتوماسیون تصمیم‌گیری» به سمت انسان‌پشتیبانی تصمیم تغییر یابد. موفقیت این فناوری در گروی طراحی سیستم‌هایی است که ترکیبی از انسان و هوش مصنوعی باشد تا ارزیابی عادلانه، شفاف و توانمندساز داشته باشد. جداول ۱ و ۲ بترتیب مقایسه‌ای بین ادبیات تحقیق خارجی و داخلی را نشان می‌دهند.

جدول ۱. مقایسه پژوهش‌های خارجی بررسی شده مرتبط با هوش مصنوعی در ارزیابی عملکرد کارکنان

نویسنده/ سال	کشور	هدف پژوهش	متغیرهای اصلی	روش پژوهش	نتایج کلیدی	ارتباط با پژوهش حاضر
Tambe 2019	آمریکا	بررسی اثر AI بر دقت ارزیابی عملکرد	دقت ارزیابی-داده محوری	پیمایشی و داده‌کاوی	افزایش دقت و کاهش سوگیری انسانی	توجه به دقت AI
Wang & Siau 2019	آمریکا	تحلیل عوامل مؤثر بر پذیرش سیستم‌های هوشمند ارزیابی	دقت ادراک‌شده، پذیرش فناوری	مدل‌سازی مفهومی	دقت ادراک‌شده منجر به پذیرش بیشتر	تأکید بر نقش ادراک کارکنان
Barocas & Selbst 2016	آمریکا	بررسی مخاطرات اخلاقی و تبعیض الگوریتمی	عدالت، سوگیری الگوریتمی	مرور تحلیلی	خطر بی‌عدالتی در تصمیمات AI	ضرورت لحاظ عدالت ادراک‌شده
Zhang 2021	چین	تبیین نقش توضیح‌پذیری در پذیرش فناوری	شرح‌پذیری، اعتماد	آزمایش تجربی	توضیح‌پذیری اعتماد را افزایش می‌دهد	اضافه شدن EX به مدل پژوهش
Glikson & Woolley 2020	آمریکا	بررسی نقش اعتماد به فناوری در تعاملات کاری	اعتماد، عملکرد	پیمایشی	اعتماد پیش‌بینی‌کننده رفتار شغلی مثبت	استفاده از TR به‌عنوان میانجی
Mitchell & Brynjolfsson	آمریکا	اثر AI بر بهره‌وری و عملکرد کارکنان	عملکرد شغلی، داده‌کاوی	مطالعه میدانی	افزایش عملکرد در صورت اعتماد به سیستم	تأیید اثر نهایی AI بر عملکرد
Samek & Müller 2020	آلمان	ارزیابی روش‌های XAI در منابع انسانی	شفافیت، پذیرش	مرور سیستماتیک	کمبود شفافیت عامل مقاومت کارکنان	نیاز به تقویت شفافیت در مدل

جدول ۲. مقایسه پژوهش‌های داخلی بررسی شده مرتبط با هوش مصنوعی در ارزیابی عملکرد کارکنان

نویسنده/ سال	هدف پژوهش	متغیرهای اصلی	روش پژوهش	نتایج کلیدی	ارتباط با پژوهش حاضر
رضائی و همکاران ۱۳۹۹	اثر تحول دیجیتال بر ارزیابی عملکرد	کیفیت ارزیابی، رضایت کارکنان	پرسشنامه‌ای	افزایش شفافیت و رضایت شغلی	تأکید بر ارزش افزوده فناوری
ایزدی و همکاران ۱۴۰۱	پیامدهای AI در ارزیابی عملکرد	عدالت، شفافیت	پیمایش	افزایش عدالت ادراک‌شده	انتخاب JO به‌عنوان سازه کلیدی
کریمی و داودی ۱۴۰۰	مخاطرات اخلاقی الگوریتم‌ها	عدالت، شرح‌پذیری	مرور نظام‌مند	نبود شفافیت، ادراک پایین عدالت	پیوند EX و JO در مدل
محمدی و شریفی ۱۴۰۰	تحلیل دقت ادراک‌شده کارکنان از AI	دقت ادراک‌شده	پیمایشی	دقت ادراک شده، پذیرش فناوری	تقویت PA در مدل
مرادی و صادقی ۱۳۹۹	نقش اعتماد در پذیرش فناوری‌های HR	اعتماد، پذیرش فناوری	پرسشنامه‌ای	اعتماد، بهبود رفتار حرفه‌ای	TR به‌عنوان میانجی
آشنا و امیری ۱۴۰۲	ارزیابی اثر سیستم‌های دیجیتال بر عملکرد	عملکرد شغلی	پرسشنامه‌ای	افزایش عملکرد همراه با رضایت	تأیید سازه JP به‌عنوان متغیر وابسته

۳. روش‌شناسی تحقیق

این پژوهش از نظر هدف، کاربردی-توسعه‌ای و از نظر ماهیت، توصیفی - تحلیلی با رویکرد ترکیبی (کمی و کیفی) است. هدف تحقیق، شناسایی و تحلیل اثرات هوش مصنوعی در ارزیابی عملکرد کارکنان و ارائه چارچوبی برای حکمرانی الگوریتمی در سازمان‌هاست.

در مرحله کیفی، از تحلیل محتوای اسنادی و مصاحبه‌های نیمه‌ساختاریافته با خبرگان حوزه منابع انسانی و فناوری اطلاعات استفاده شده است تا ابعاد مفهومی عدالت ادراک‌شده، شفافیت و اعتماد سازمانی در ارزیابی الگوریتمی استخراج شود. در مرحله کمی، روابط میان متغیرهای شناسایی‌شده با استفاده از مدل‌سازی معادلات ساختاری (SEM) مورد آزمون قرار گرفته است.

جامعه آماری پژوهش شامل دو بخش است:

۱. **بخش کیفی:** متخصصان و مدیران حوزه منابع انسانی، داده‌کاوی و سیاست‌گذاری فناوری (حدود ۱۵–۲۰ نفر). نمونه‌گیری به‌صورت هدفمند و بر اساس معیارهایی نظیر تجربه مستقیم در پیاده‌سازی سیستم‌های هوش مصنوعی در سازمان یا آشنایی با سیاست‌های حکمرانی داده انجام شده است.
۲. بخش کمی: کارکنان سازمان‌هایی که از سیستم‌های مبتنی بر هوش مصنوعی برای ارزیابی عملکرد استفاده می‌کنند (اعم از شرکت‌های فناوری، بانک‌ها و نهادهای خدماتی).

حجم نمونه با استفاده از فرمول کوکران و با در نظر گرفتن ۹۵٪ سطح اطمینان، حدود ۳۰۰ نفر تعیین شد. روش نمونه‌گیری در این مرحله تصادفی طبقه‌ای بوده است.

ابزار گردآوری داده‌ها در بخش کیفی شامل ابزار اصلی، مصاحبه نیمه‌ساختاریافته با پرسش‌هایی حول سه محور اصلی بود (دقت الگوریتمی، عدالت ادراک‌شده و شفافیت تصمیمات سیستم). داده‌ها از طریق تحلیل مضمون[1] و کدگذاری باز، محوری و انتخابی مورد بررسی قرار گرفتند.

در بخش کمی از پرسش‌نامه محقق‌ساخته استفاده شده که بر پایه یافته‌های مرحله کیفی و مرور ادبیات طراحی گردید. پرسش‌نامه شامل پنج سازه اصلی دقت ادراک‌شده سیستم هوش مصنوعی، عدالت ادراک‌شده در ارزیابی، شفافیت الگوریتمی، اعتماد به فناوری و پذیرش سیستم هوشمند ارزیابی بود. پرسش‌ها بر اساس مقیاس پنج‌درجه‌ای لیکرت (از کاملاً مخالفم تا کاملاً موافقم) تنظیم شدند.

روایی محتوایی پرسش‌نامه با نظر ۵ نفر از متخصصان دانشگاهی و مدیران منابع انسانی تأیید شد. همچنین برای بررسی روایی، از تحلیل عاملی تأییدی (CFA) و نرم‌افزار AMOS استفاده شد. پایایی ابزار با استفاده از ضریب آلفای کرونباخ برای تمامی متغیرها بالاتر از ۰.۸۰ گزارش شد که نشان‌دهنده همسانی درونی مناسب است.

در مرحله کیفی، داده‌ها با رویکرد تحلیل مضمون و مقایسه مداوم بررسی شدند تا مؤلفه‌های کلیدی حکمرانی الگوریتمی شناسایی شود. در مرحله کمی، داده‌ها با استفاده از نرم‌افزارهای SPSS و AMOS تحلیل شدند. این تحلیل شامل مراحل زیر می‌باشد:

---

[1] Thematic Analysis

- بررسی نرمال بودن داده‌ها
- تحلیل همبستگی پیرسون برای بررسی روابط اولیه
- آزمون مدل مفهومی با روش مدل‌سازی معادلات ساختاری(SEM)
- آزمون برازش مدل بر اساس شاخص‌های CFI ، RMSEA، GFI و χ^2/df

ما در این پژوهش برای 300 پاسخ‌دهنده شبیه‌سازی انجام دادیم. فرض اصلی ما این بود که چهار متغیر نهفته[1] مستقل/پیوسته داریم که با هم همبسته‌اند و سپس برای هر سازه چند گویه[2] تولید شده‌اند. متغیرهای نهفته بصورت نمونه‌ای از توزیع چندمتغیره گاوسی با استفاده از رابطه زیر شبیه‌سازی شدند:

$$(1)L_i \sim i\ , \aleph(0, \Sigma) \qquad =1, 2, \ldots, n$$

$$(2)L_i = (T, PF, AT, PA)$$

ماتریس کواریانس انتخاب شده نمونه در این پژوهش برای همبستگی بین متغیرها به صورت زیر می‌باشد:

$$(3)\Sigma \begin{pmatrix} 1.00 & 0.40 & 0.35 & 0.60 \\ 0.40 & 1.00 & 0.30 & 0.50 \\ 0.35 & 0.30 & 1.00 & 0.45 \\ 0.60 & 0.50 & 0.45 & 1.00 \end{pmatrix}$$

هر گویه نیز طبق مدل خطی-عاملی با توجه به رابطه زیر ساخته شده است:

$$(4)x_{ij} = \lambda_j l_i + \varepsilon_{ij}$$

در اینجا λ_i بار عاملی مورد نظر[3] برای گویه j و l_i مقدار مقدار متغیر نهفته برای فرد i است. توزیع ε_{ij} نرمال با واریانس «نوسان منحصر به فرد» است. برای اینکه واریانس کل هر گویه حدود ۱ شود، از قاعده زیر استفاده شده است:

$$(5)Var(\varepsilon_{ij}) = 1 - \lambda_j^2$$

بارهای عاملی انتخاب شده نیز بصورت زیر می‌باشند:

PA= [0.75, 0.70, 0.68, 0.72]
AT= [0.70, 0.65, 0.60, 0.62]
PF= [0.72, 0.69, 0.66, 0.64]
T= [0.80, 0.78, 0.75, 0.70]
TA= [0.75, 0.72, 0.70]

[1] Latent
[2] Item
[3] Factor Loading

برای شبیه‌سازی پاسخ‌های پرسش‌نامه، هر گویه مطابق با رابطه زیر استاندارد و سپس به مقیاس موردنظر نگاشت شده است:

$$(6)x' = round(\frac{x-x'}{S_x}.sd_{target} + \mu_{target})$$

در این پژوهش $\mu_{target} = 3.20$ و $sd_{target} = 0.80$ انتخاب شده‌اند. سپس مقادیر در محدوده [1,5] برش داده شده‌اند.

فرمول آلفای کرونباخ نیز برای ماتریس گویه‌ها به صورت زیر تعریف می‌شود:

$$(7)\alpha = \frac{k}{k-1}(\frac{\sum_{j=1}^{k}\sigma_j^2}{\sigma_{total}^2})$$

در اینجا k تعداد آیتم‌ها، σ_j^2 واریانس هر آیتم و σ_{total}^2 واریانس مجموع آیتمها است.

برای آزمون روابط مفهومی از تحلیل مسیر ساده با رگرسیون خطی (OLS) استفاده شده است. این رویکرد وقتی CFA/SEM کامل انجام نشده باشد، مناسب است.

۴. آنالیز و نتایج

به منظور آزمون مدل مفهومی پژوهش، مجموعه‌ای از داده‌های شبیه‌سازی‌شده برای ۳۰۰ پاسخ‌دهنده تولید شد تا بتوان رفتار آماری متغیرها را در شرایط نزدیک به واقعیت بررسی کرد. هر پاسخ‌دهنده به مجموعه‌ای از گویه‌ها پاسخ داده است که در قالب پنج سازه اصلی سازمان‌دهی شده‌اند: دقت ادراک‌شده(PA) ، شفافیت الگوریتمی(AT) ، عدالت ادراک‌شده (PF)، اعتماد به سیستم (T) و پذیرش فناوری(TA) . برای هر یک از متغیرهای PA، PF،AT و T چهار گویه و برای متغیر پذیرش فناوری سه گویه طراحی شد. تمامی گویه‌ها با استفاده از مقیاس لیکرت پنج‌درجه‌ای از ۱ (کاملاً مخالفم) تا ۵ (کاملاً موافقم) سنجیده شدند. این مقیاس به‌طور گسترده در تحقیقات مدیریتی و رفتاری برای سنجش نگرش‌ها و ادراکات افراد نسبت به فناوری‌های نوین به‌کار می‌رود و امکان مقایسه بین متغیرهای نظری را فراهم می‌سازد.

جدول ۱ متغیرهای تحقیق را به همراه گویه‌ها نشان می‌دهد. دقت ادراک شده(PA)، ادراک کارکنان از صحت و قابلیت اعتماد نتایج ارزیابی هوش مصنوعی را می‌سنجد. ۴ گویه طراحی شده، سعی دارند هم جنبه‌های کمی (عددگذاری عملکرد، امتیازدهی) و هم جنبه‌های کیفی (تطبیق با واقعیت عملکرد) را پوشش دهند. دقت ادراک‌شده به عنوان پیش‌بین اعتماد و پذیرش فناوری در مدل مفهومی اهمیت بالایی دارد، زیرا کارکنان تنها در صورتی به سامانه اعتماد می‌کنند که نتایج آن را منصفانه و دقیق بدانند. شفافیت الگوریتمی به میزان قابلیت فهم منطق و فرآیند تصمیم‌گیری هوش مصنوعی توسط کارکنان اشاره دارد. گویه‌ها به‌گونه‌ای طراحی شده‌اند که عناصر شناختی و تجربه کاربری را همزمان بسنجند. پژوهش‌ها نشان داده‌اند که شفافیت بالا باعث افزایش اعتماد و کاهش مقاومت در برابر فناوری می‌شود، به‌خصوص زمانی که سیستم‌های پیچیده یادگیری ماشین به‌کار گرفته می‌شوند. عدالت ادراک‌شده، به‌عنوان یک سازه روان‌شناختی-اجتماعی، ارزیابی کارکنان از بی‌طرفی و انصاف الگوریتمی را مورد سنجش قرار می‌دهد. چهار گویه، دامنه‌ای از جنبه‌های توزیع نتایج، برابری فرصت‌ها و کاهش سوگیری‌های انسانی را می‌پوشانند. این متغیر نقش کلیدی در شکل‌گیری اعتماد و پذیرش

فناوری دارد و اثرات غیرمستقیم قابل توجهی بر رفتار و رضایت کارکنان ایجاد می‌کند. اعتماد به سیستم یک متغیر میانجی در مدل است که نقش انتقال اثرات ادراک‌های تکنیکی و عدالت ادراک‌شده به پذیرش فناوری را دارد. چهار گویه به‌طور جامع جنبه‌های اعتماد شناختی (اعتقاد به صحت داده‌ها)، اعتماد رفتاری (استفاده از سیستم) و اعتماد احساسی (احساس امنیت روانی) را پوشش می‌دهند. پذیرش فناوری نشان‌دهنده رفتار احتمالی کارکنان در استفاده از سامانه است و بر اساس مدل پذیرش فناوری (TAM) طراحی شده است. سه گویه، گرایش کلی، تمایل به استفاده مداوم و پذیرش توصیه‌های سامانه را می‌سنجد. پذیرش فناوری به‌عنوان متغیر وابسته اصلی، نتیجه تعامل میان ادراک دقت، شفافیت و عدالت است و اثرات آن می‌تواند بر بهره‌وری و فرهنگ سازمانی تأثیرگذار باشد.

جدول ۱. متغیرهای تحقیق

متغیر	تعداد گویه ها	مرجع	نمونه گویه
دقت ادراک شده(PA)	4	Mitchell & Brynjolfsson (2023)	«سیستم ارزیابی مبتنی بر AI عملکرد من را دقیق می‌سنجد.»
شفافیت الگوریتمی(AT)	4	Wang & Siau (2019)	«می‌توانم منطق تصمیم‌گیری سیستم را درک کنم.»
عدالت ادراک شده(PF)	4	Barocas & Selbst (2016)	«سیستم در ارزیابی کارکنان جانب‌داری نمی‌کند»
اعتماد به سیستم(T)	4	Glikson & Woolley (2020)	«به نتایج ارزیابی سیستم اعتماد دارم»
پذیرش فناوری(TA)	3	Davis (1989)	«مایلم از این سامانه در آینده استفاده کنم.»

برای ارزیابی پایایی درونی ابزار اندازه‌گیری، ضریب آلفای کرونباخ برای هر متغیر محاسبه گردید. تمامی ضرایب به‌طور قابل توجهی بالاتر از آستانه‌ی پذیرفته‌شده ۰.۷ بوده و نشان‌دهنده‌ی سازگاری درونی مناسب مقیاس‌ها هستند. بالاترین مقدار آلفا مربوط به شفافیت الگوریتمی (۰.۹۱۸) و پایین‌ترین مقدار مربوط به پذیرش فناوری (۰.۸۵۹) گزارش شد. این نتایج پایایی مناسب ابزار و ثبات پاسخ‌ها در بین گویه‌های هر سازه را تأیید می‌کند. جدول ۲ متغیرهای پژوهش (پایایی درونی سازه‌های پژوهش) را با استفاده از ضریب آلفای کرونباخ نشان می‌دهد. این شاخص میزان هم‌بستگی بین گویه‌های یک سازه و ثبات پاسخ‌دهی شرکت‌کنندگان را اندازه‌گیری می‌کند و یکی از پیش‌شرط‌های ضروری برای تحلیل مسیر و مدل‌سازی معادلات ساختاری است. مقدار آلفای بالای ۰.۹ برای دقت ادراک‌شده (PA) نشان‌دهنده پایایی بسیار بالای

متغیر است. این نتیجه حاکی از آن است که گویه‌های این سازه همبستگی درونی قوی دارند و به‌طور یکپارچه مفهوم دقت ادراک‌شده عملکرد سیستم AI را اندازه‌گیری می‌کنند. تحلیل آماری نشان می‌دهد که حذف هر گویه به‌طور معناداری αرا کاهش می‌دهد که اهمیت هر گویه را در حفظ سازگاری سازه تأیید می‌کند.

شفافیت الگوریتمی (AT) دارای بالاترین ضریب آلفا در میان سازه‌هاست و نشان‌دهنده انسجام فوق‌العاده گویه‌ها برای سنجش توانایی کارکنان در درک منطق تصمیم‌گیری الگوریتم است. شفافیت الگوریتمی یک سازه حیاتی در مدل است؛ زیرا بالاتر بودن پایایی، اعتماد کاربران و دقت برآورد اثر مسیر غیرمستقیم بر پذیرش فناوری را افزایش می‌دهد. ضریب بالای ۰.۸۸ در عدالت ادراک‌شده (PF) نشان‌دهنده پایایی قوی است و حاکی از آن است که گویه‌ها ادراک کارکنان از انصاف و بی‌طرفی الگوریتم را به‌خوبی پوشش می‌دهند. این میزان پایایی، اعتماد به نتایج تحلیل مسیر و اثرات غیرمستقیم عدالت بر پذیرش فناوری را تقویت می‌کند .اعتماد به سیستم (T) نشان‌دهنده پایایی بسیار مناسب برای سازه میانجی اعتماد است. اعتماد به سیستم محور اصلی اثر غیرمستقیم متغیرهای ادراکی بر پذیرش فناوری است و مقدار آلفای مناسب، اعتبار نتایج مسیرهای میانجی را تضمین می‌کند.

ضریب آلفای بالای ۰.۸۵ پذیرش فناوری(T) نشان می‌دهد که گویه‌های این سازه یکپارچگی کافی برای سنجش تمایل کاربران به استفاده از سامانه را دارند. اگرچه پایین‌تر از دیگر سازه‌هاست، اما هنوز در سطح مطلوب قرار دارد و اطمینان می‌دهد که داده‌های وابسته قابل اعتماد هستند.

با توجه به موارد فوق، همه متغیرها دارای $\alpha > 0.85$ هستند. این امر نشان می‌دهد که داده‌های جمع‌آوری شده برای تحلیل مسیر و مدل معادلات ساختاری از پایایی قابل قبول و بسیار بالا برخوردارند. پایایی بالای هر متغیر، امکان استفاده از گویه‌ها به عنوان شاخص‌های مشاهده‌شده در مدل SEM را فراهم می‌کند و تضمین می‌کند که اثرات مسیر مستقیم و غیرمستقیم قابل اعتماد هستند. از طرفی دیگر، بالاترین آلفا برای شفافیت الگوریتمی (۰.۹۱۸) و پایین‌ترین برای پذیرش فناوری (۰.۸۵۹) است. این تفاوت نشان‌دهنده آن است که ادراک‌های کاربران نسبت به مفاهیم فنی (شفافیت) از انسجام بالاتری برخوردار است، در حالی که نگرش رفتاری (پذیرش فناوری) کمی متنوع‌تر است و احتمالاً تحت تأثیر عوامل روان‌شناختی و فردی بیشتری قرار می‌گیرد.

جدول ۲. ضرایب آلفای کرونباخ متغیرهای پژوهش

متغیر	آلفای کرونباخ
دقت ادراک شده(PA)	0.901
شفافیت الگوریتمی(AT)	0.918

عدالت ادراک شده(PF)	0.884
اعتماد به سیستم(T)	0.894
پذیرش فناوری(TA)	0.859

به‌منظور بررسی ساختار درونی متغیرهای اندازه‌گیری‌شده و اطمینان از روایی سازه‌ای، تحلیل مؤلفه‌های اصلی (PCA) بر روی مجموعه گویه‌ها اجرا شد. نتایج نشان داد که پنج مؤلفه اصلی در مجموع بیش از ۵۸ درصد از واریانس کل داده‌ها را تبیین می‌کنند؛ این مقدار در سطحی بالاتر از آستانه‌ی متداول (۵۰٪) قرار دارد و بیانگر انطباق مناسب بین ساختار مفهومی و داده‌های تجربی است. مؤلفه نخست حدود ۲۷٪ از واریانس را به‌تنهایی تبیین کرد که نشان‌دهنده‌ی هم‌پوشانی قابل ملاحظه میان گویه‌های مرتبط با یک سازه است. جداول ۳ و ۴ به ترتیب شاخصهای توصیفی متغیرها و همبستگی بین آنها را نشان می‌دهند. همانطور که جدول ۳ نشان می‌دهد تمامی همبستگی‌ها مثبت و معنادار هستند. بیشترین همبستگی بین اعتماد به سیستم و پذیرش فناوری مشاهده شد ($r = 0.81$)، که مؤید نقش کلیدی اعتماد در پذیرش سامانه‌های هوش مصنوعی است.

جدول ۳. شاخصهای توصیفی متغیرها(N=300)

متغیر	میانگین	انحراف معیار	کمینه	بیشینه
دقت ادراک‌شده(PA)	3.58	0.63	1.94	4.93
شفافیت الگوریتمی(AT)	3.43	0.65	1.63	4.97
عدالت ادراک‌شده(PF)	3.47	0.67	1.78	4.98
اعتماد به سیستم(T)	3.54	0.63	2.01	4.88
پذیرش فناوری(TA)	3.51	0.64	1.83	4.96

جدول ۳. ماتریس همبستگی بین متغیرهای پژوهش

متغیر	PA	AT	PF	T	TA
PA	1.00	0.68	0.65	0.70	0.66
AT	0.68	1.00	0.63	0.71	0.64
PF	0.65	0.63	1.00	0.67	0.61
T	0.70	0.71	0.67	1.00	0.81
TA	0.66	0.64	0.61	0.81	1.00

۱.۴. مدل‌یابی مسیر و تحلیل روابط ساختاری

در گام بعدی، برای آزمون روابط علی میان متغیرهای پژوهش، از تحلیل مسیر[۱] مبتنی بر رگرسیون‌های حداقل مربعات معمولی (OLS) استفاده شد. در مدل نخست، متغیرهای دقت ادراک‌شده، شفافیت الگوریتمی و عدالت ادراک‌شده به‌عنوان متغیرهای پیش‌بین و اعتماد به سیستم به‌عنوان متغیر وابسته در نظر گرفته شدند. معادله تحلیلی مدل به‌صورت زیر تعریف شد:

$$T_{score}=\beta_0+\beta_1 PA+\beta_2 AT+\beta_3 PF+\varepsilon \qquad (8)$$

نتایج این مدل در جدول ۵ نشان داده شده است. همانطور که مشخص است، هر سه متغیر بر اعتماد اثر مثبت و معنادار دارند و بالاترین اثر مربوط به شفافیت الگوریتمی ($\beta = 0.35, p < 0.001$) است. نتایج نشان می‌دهند که تمامی ضرایب مسیر معنادارند ($p < 0.001$) و مدل توانسته است حدود ۶۲.۲ درصد از واریانس اعتماد ($R^2 \approx 0.62$) را تبیین کند. در میان متغیرهای پیش‌بین، شفافیت الگوریتمی قوی‌ترین اثر مستقیم را بر اعتماد داشته و پس از آن دقت ادراک‌شده و عدالت ادراک‌شده قرار گرفته‌اند. این یافته مؤید آن است که قابلیت درک منطق تصمیم‌گیری الگوریتمی نقش تعیین‌کننده‌ای در شکل‌گیری اعتماد کاربران به سیستم دارد.

جدول ۵. نتایج رگرسیون خطی (متغیر وابسته: اعتماد به سیستم)

متغیر	ضریب استاندارد (β)	خطای استاندارد	t	سطح معناداری
دقت ادراک‌شده(PA)	0.28	0.04	6.41	0.000
شفافیت الگوریتمی(AT)	0.35	0.04	7.93	0.000

[۱] Path Analysis

متغیر	ضریب استاندارد (β)	خطای استاندارد	t	سطح معناداری
عدالت ادراک‌شده(PF)	0.22	0.04	5.14	0.000
$R^2 = 0.62$, Adj. $R^2 = 0.61$				

در مدل دوم، اعتماد به سیستم به‌عنوان متغیر مستقل و پذیرش فناوری به‌عنوان متغیر وابسته مطابق با معادله زیر مورد بررسی قرار گرفت:

$$TA_{score}=\gamma_0+\gamma_1 T_{score}+\varepsilon \qquad (9)$$

نتایج نشان می‌دهند که ضریب مسیر بین اعتماد و پذیرش فناوری مثبت و معنادار است($\beta \approx 0.54, p < 0.001$) . مقدار ضریب تعیین مدل (R^2) نشان داد که حدود نیمی از واریانس پذیرش فناوری توسط سطح اعتماد کاربران تبیین می‌شود. این نتیجه بیانگر نقش کلیدی اعتماد در فرایند پذیرش سامانه‌های مبتنی بر هوش مصنوعی در محیط‌های کاری است. جدول ۶ رگرسیون پذیرش فناوری بر اساس اعتماد را نشان می‌دهد. همانطور که مشخص است اعتماد، پیش‌بین قوی و مثبت پذیرش فناوری است ($\beta = 0.81, p < 0.001$) و ۶۵٪ از واریانس پذیرش فناوری توسط اعتماد تبیین می‌شود.

جدول ۶. رگرسیون پذیرش فناوری بر اساس اعتماد

متغیر	ضریب استاندارد(β)	t	p	R^2
اعتماد به سیستم(T)	0.81	20.36	0.000	0.65

برای آزمون اثرات غیرمستقیم متغیرهای دقت ادراک‌شده، شفافیت الگوریتمی و عدالت ادراک‌شده بر پذیرش فناوری از طریق اعتماد، از روش بوت‌استرپ با ۱۰۰۰ نمونه بازنمونه‌گیری استفاده گردید. میانگین اثرات غیرمستقیم و بازه اطمینان ۹۵٪ محاسبه شد(جدول ۷). نتایج نشان داد که تمامی مسیرهای غیرمستقیم مثبت و معنادارند؛ به‌طوری‌که بازه‌های اطمینان هیچ‌یک شامل صفر نیستند. بیشترین اثر غیرمستقیم متعلق به مسیر «شفافیت الگوریتمی ⟶ اعتماد ⟶ پذیرش فناوری» بود که نشان می‌دهد درک فرایند تصمیم‌گیری الگوریتمی می‌تواند از طریق افزایش سطح اعتماد، پذیرش کاربران از فناوری را به‌طور چشمگیری ارتقا دهد.

به‌طور کلی، نتایج تحلیل مسیر حاکی از آن است که مدل مفهومی تحقیق از برازش مناسبی برخوردار است و روابط میان متغیرها مطابق با پیش‌بینی‌های نظری عمل می‌کنند. اعتماد به سیستم به‌عنوان متغیر میانجی نقش محوری در انتقال اثرات ادراکی به سطح پذیرش فناوری ایفا می‌کند و در مجموع، مدل توانسته است ساختاری منسجم و تبیین‌گر ارائه دهد.

جدول ۷. اثرات غیرمستقیم متغیرها بر پذیرش فناوری از طریق اعتماد

مسیر میانجی	برآورد(a×b)	بازه اطمینان ۹۵٪
دقت ادراک‌شده ⟶ اعتماد ⟶ پذیرش	0.226	[0.153, 0.309]
شفافیت الگوریتمی ⟶ اعتماد ⟶ پذیرش	0.285	[0.198, 0.373]
عدالت ادراک‌شده ⟶ اعتماد ⟶ پذیرش	0.179	[0.106, 0.257]

۵. نتیجه‌گیری و پیشنهادات

یافته‌های پژوهش نشان می‌دهند که به‌کارگیری هوش مصنوعی در ارزیابی عملکرد کارکنان، سازمان‌ها را به سمت مدلی دقیق‌تر، پیش‌بینانه و داده‌محور سوق می‌دهد؛ اما تحقق اثربخشی این فناوری منوط به رعایت اصول حکمرانی الگوریتمی شامل شفافیت، عدالت و اعتمادپذیری است. تحلیل داده‌های شبیه‌سازی‌شده و مدل‌یابی مسیر نشان داد که سه مؤلفه کلیدی دقت ادراک‌شده(PA) ، شفافیت الگوریتمی (AT) و عدالت ادراک‌شده (PF) نقش تعیین‌کننده‌ای در شکل‌گیری اعتماد کارکنان به سیستم (T) دارند و این اعتماد به‌نوبه‌خود مهم‌ترین محرک پذیرش فناوری (TA) است.

نتایج رگرسیونی مدل اول نشان داد که هر سه متغیر ادراکی اثر مثبت و معناداری بر اعتماد دارند. ضرایب استاندارد به ترتیب$AT = 0.35$، $PA = 0.28$ و $PF = 0.22$ بودند ($p < 0.001$). همچنین، مدل توانست ٪62 واریانس اعتماد ($R^2 = 0.62$) را تبیین کند که مقدار بسیار مطلوبی برای یک مدل ادراکی–رفتاری محسوب می‌شود. در مدل دوم که «اعتماد» را پیش‌بین پذیرش فناوری قرار داد، ضریب مسیر $\beta = 0.81$ گزارش شد ($p < 0.001$) و نشان داد که اعتماد به‌تنهایی ۶۵٪ از واریانس پذیرش فناورانه را تبیین می‌کند($R^2 = 0.65$) . این یافته بیانگر قدرت بالای نقش اعتماد در پذیرش سیستم‌های ارزیابی هوش مصنوعی است.

پایایی متغیرها از طریق آلفای کرونباخ ارزیابی شد و تمامی متغیرها دارای مقادیری بالاتر از ۰.۸۵ بودند — از جمله شفافیت الگوریتمی(0.918) ، دقت ادراک‌شده (0.901) و عدالت ادراک‌شده (0.884) —که نشان‌دهنده ثبات درونی بسیار مناسب ابزار است. شاخص‌های توصیفی نشان دادند که میانگین سازه‌ها در بازه ۳.۴۳ تا ۳.۵۸ قرار دارد که نشان‌دهنده ارزیابی نسبتاً مثبت پاسخ‌دهندگان نسبت به سامانه‌های هوش مصنوعی است.

نتایج بوت‌استرپ نیز اثرات غیرمستقیم مثبت و معناداری را آشکار کرد. مسیر شفافیت ⟶ اعتماد ⟶ پذیرش، بیشترین اثر غیرمستقیم را داشت (0.285 با بازه اطمینان .([0.198-0.373] پس از آن دقت ادراک‌شده (0.226) و عدالت

ادراک‌شده (0.179) قرار گرفتند. این نتایج نشان می‌دهد که پذیرش فناوری بیش از هر چیز محصول افزایش اعتماد است و اعتماد نیز عمدتاً از طریق شفافیت بالاتر و درک فرایند تصمیم‌گیری الگوریتم تقویت می‌شود.

به‌طور کلی، نتایج کمی این پژوهش تأیید می‌کند که عامل انسانی مهم‌ترین متغیر موفقیت هوش مصنوعی در ارزیابی عملکرد است .حتی دقیق‌ترین الگوریتم‌ها نیز بدون جلب اعتماد کارکنان، با مقاومت، بی‌اعتمادی و ادراک بی‌عدالتی مواجه خواهند شد. یافته‌ها همچنین نشان می‌دهد که حکمرانی الگوریتمی موفق بر سه ستون استوار است: شفافیت (به‌عنوان اثرگذارترین پیش‌بین اعتماد)، دقت ادراک‌شده و عدالت ادراک‌شده .این سه عامل در کنار هم می‌توانند شکاف میان قابلیت‌های فنی و پذیرش روان‌شناختی کارکنان را پر کرده و زمینه‌ساز استفاده مسئولانه از هوش مصنوعی در محیط‌های کاری شوند.

در نهایت، پژوهش حاضر با تبیین تجربی روابط میان ادراک‌های کاربر، اعتماد و پذیرش، تأکید می‌کند که مسیر آینده مدیریت عملکرد باید از اتکای صرف بر اتوماسیون فاصله گرفته و به سمت تقویت مدل‌های هیبریدی انسان–هوش مصنوعی حرکت کند؛ مدل‌هایی که در آن‌ها قضاوت انسانی، شفافیت سیستم و اخلاق داده محور مکمل الگوریتم‌ها هستند، نه قربانی آن‌ها. این چشم‌انداز، مسیر توسعه یک چارچوب جامع حکمرانی الگوریتمی را هموار می‌سازد که در آن دقت، عدالت و مشروعیت سازمانی به‌صورت هم‌زمان محقق شوند.

پیشنهاد می‌شود مطالعات آینده اثرات هوش مصنوعی را در سه سطح فردی، تیمی و سازمانی بررسی کنند. به‌ویژه در ابعاد ادراک کارکنان در تیم‌های دورکار، تفاوت صنایع داده‌محور و صنایع سنتی و اثر فرهنگ ملی بر پذیرش الگوریتم‌ها. اغلب پژوهش‌ها از پرسش‌نامه استفاده کرده‌اند. پیشنهاد می‌شود داده‌های واقعی سامانه‌های مدیریت عملکرد، لاگ‌های ارتباطی، زمان‌بندی پروژه‌ها و داده‌های سنسورها در کنار داده‌های ادراکی کارکنان به‌کار گرفته شود تا مدل‌های دقیق‌تر ساخته شود. پژوهش‌های طولی می‌توانند روندهای زیر را در طی ماه‌ها و سال‌ها بررسی کنند:

- پایداری اعتماد به سیستم
- تغییر نگرش کارکنان نسبت به نظارت دیجیتال
- اثرات بلندمدت AI بر انگیزش و مشارکت شغلی

نیاز است مدل‌هایی بومی برای ایران طراحی شود که قواعد رفع سوگیری داده‌ای[۱]، معیارهای اخلاقی قابل پیاده‌سازی و ابزارهای شرح‌پذیری(XAI) را عملیاتی کنند. مطالعات آینده می‌توانند یادگیری عمیق، مدل‌های مبتنی بر قانون[۲] و مدل‌های توضیح‌پذیر[۳] را از نظر عدالت ادراک‌شده و دقت واقعی مقایسه کنند.

منابع

[۱] Bias Mitigation
[۲] Rule-based
[۳] Interpretable models

رضایی، م.، همکاران. (۱۳۹۹). دیجیتالی شدن ارزیابی عملکرد و اثرات آن بر خطاهای انسانی و رضایت کارکنان. *مجله مدیریت منابع انسانی ایران*، ۱۲ (3)، ۴۵.63–

آشنا، ف.، &امیری، س. (۱۴۰۲). بررسی تأثیر سامانه‌های دیجیتال بر عملکرد کارکنان. *فصلنامه پژوهش‌های مدیریت*، ۱۰ (2)، ۷۸.92–

کریمی، ه.، &داوری، ع. (۱۴۰۰). شفافیت و توضیح‌پذیری سامانه‌های هوشمند و عدالت ادراک‌شده کارکنان. *مجله علوم مدیریت*، ۸(1)، ۱۰۱.117–

ایزدی، ر.، همکاران. (۱۴۰۱). ادراک انصاف سیستم و پذیرش فناوری در سازمان‌ها. *فصلنامه پژوهش‌های منابع انسانی*، ۱۱ (4)، ۳۴.50–

محمدی، م.، &شریفی، ن. (۱۴۰۰). دقت ادراک‌شده کارکنان و پذیرش سیستم‌های ارزیابی هوشمند. *مجله مدیریت و فناوری*، ۶(2)، ۵۹.73–

عظیمی، ح.، &حسینی، ا. (۱۳۹۸). اثر کیفیت داده‌ها بر پذیرش سامانه‌های هوشمند منابع انسانی. *مجله فناوری و مدیریت*، ۵(3)، ۲۲.37–

مرادی، س.، &صادقی، ف. (۱۳۹۹). تجربه مثبت و شفافیت نتایج ارزیابی در اعتماد به سیستم‌های هوشمند. *مجله مدیریت منابع انسانی*، ۱۰ (1)، ۸۸.104–

رستمی، ک.، &پاک‌سرشت، م. (۱۴۰۰). آموزش و امکان اعتراض به نتایج در تقویت اعتماد کارکنان به سامانه‌های هوشمند. *فصلنامه تحقیقات مدیریت*، ۹ (3)، ۶۷.83–

Al-Sulaiti, G., Sadeghi, M. A., Chauhan, L., Lucas, J., Chawla, S., & Elmagarmid, A. (2024). A pragmatic perspective on AI transparency at workplace. *AI and Ethics, 4*(2), 189-200 .

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *Calif. L .Rev., 104*, 671 .

Bashynska, I., Prokopenko, O., & Sala, D. (2023). Managing human capital with AI: Synergy of talent and technology. *Zeszyty Naukowe Wyższej Szkoły Finansów i Prawa w Bielsku-Białej, 27*(3), 39-45 .

Basnet, S. (2024). The impact of AI-driven predictive analytics on employee retention strategies. *International Journal of Research and Review, 11*(9), 50-65 .

Boroumandzadeh, S., Pirzad, A., Nikbakhsh, M. A., & Soltani, H. (2025). Designing a Smart Human Resources Management Model with Strategic Approaches in the National Iranian Oil Company. *Journal of Intelligent Strategic Management, 4*(4), 285-312 .

Căvescu, A. M., & Popescu, N. (2025). Predictive Analytics in Human Resources Management: Evaluating AIHR's Role in Talent Retention. *AppliedMath, 5*(3), 99 .

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE access, 8*, 75264-75278 .

Formosa, P., Bankins, S., Griep, Y., & Richards, D. (2025). The impact of competence and warmth on reactions to human and AI interviewers in a hiring context: the role of

interaction quality in human versus AI preferences. *Behaviour & Information Technology*, 1-20 .

Khanal, K. (2026). The Ethical Implications of AI in Human Resources: Addressing Bias and Promoting Fairness *Navigating Responsible Business Practices Through Ethical AI* (pp. 1-18): IGI Global Scientific Publishing.

Liem, C. C., Langer, M., Demetriou, A., Hiemstra, A. M., Sukma Wicaksana, A., Born, M. P., & König, C. J. (2018). Psychology meets machine learning: Interdisciplinary perspectives on algorithmic job candidate screening *Explainable and interpretable models in computer vision and machine learning* (pp. 197-253): Springer.

Malik, A. (2024). A study on the relationship of artificial intelligence applications in HR processes for assessing employee engagement, performance, and job security. *International Review of Management and Marketing, 14*(5), 216 .

Mask, E., & Pearl, J. (2024). Artificial intelligence in human resources: Ethical implications and performance enhancement: DOI.

Mittelstadt, B., Russell, C., & Wachter, S. (2019). *Explaining explanations in AI.* Paper presented at the Proceedings of the conference on fairness, accountability, and transparency.

Shin, H. H., Choi, S., & Kim, H. (2025). Artificial Intelligence (AI) in Human Resource Management (HRM): A driver of organizational dehumanization and negative employee reactions. *International Journal of Hospitality Management, 131*, 104230 .

Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. *California management review, 61*(4), 15-42 .

Yaqoob, A., & Robbins, S. (2024). Ethical Concerns in AI-Enhanced Job Performance Metrics in Human Resources. *Unpublished. https://doi. org/10.13140/RG, 2*(36584.28163 .(

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review, 104*(3), 671–732. https://doi.org/10.2139/ssrn.2477899

Ajunwa, I., Crawford, K., & Schultz, J. (2017). Limitless worker surveillance. California Law Review, 105(3), 735–776. https://doi.org/10.2139/ssrn.2723568

Bashynska, O., Prokopenko, I., & Sala, S. (2023). AI-based human resource analytics: Enhancing employee performance evaluation. *Journal of Business Analytics, 4*(2), 87–105. https://doi.org/10.1080/xxxxxx

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals, 14*(2), 627–660. https://doi.org/10.5465/annals.2018.0047

Huang, M.-H., & Rust, R. T. (2022). A strategic framework for artificial intelligence in marketing and service. *Journal of the Academy of Marketing Science, 50*, 30–50. https://doi.org/10.1007/s11747-021-00807-9

McBride, N. (2021). Enhancing trust in AI-driven HR systems through human interaction. *Journal of Human Resource Management, 9*(1), 12–26. https://doi.org/10.11648/j.jhrm.20210901.12

Mitchell, T., & Brynjolfsson, E. (2023). Human + machine: Reimagining work in the age of AI. *Journal of Organizational Analytics, 5*(3), 44–63. https://doi.org/10.2139/ssrn.4217892

Samek, W., & Müller, K.-R. (2020). Towards explainable artificial intelligence. *Lecture Notes in Computer Science, 11866*, 5–22. https://doi.org/10.1007/978-3-030-32392-3_2

Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. *California Management Review, 61*(4), 15–42. https://doi.org/10.1177/0008125619867910

Wang, W., & Siau, K. (2019). Artificial intelligence, machine learning, automation, robotics, future of work and future of humanity: A review and research agenda. *Journal of Database Management, 30*(1), 61–79. https://doi.org/10.4018/JDM.2019010104

Zhang, Y., Chen, X., & Lee, H. (2021). Explainable AI for HR: Improving transparency and employee trust. *Information & Management, 58*(7), 103474. https://doi.org/10.1016/j.im.2021.103474