پژوهش‌های نوین در تصمیم‌گیری
دوره ۱۰، شماره ۳، پاییز ۱٤۰٤، صص ۸۷-۱۱٦
نوع مقاله: پژوهشی

ارائه مدلی ترکیبی جهت پیش‌بینی قصد خرید مصرف‌کنندگان مبتنی بر یادگیری ماشین بر اساس نحوه بازخورد در رسانه‌های اجتماعی

منیره حسینی *[۱]، محدثه کرکبودی [۲]

۱. دانشیار، گروه فناوری اطلاعات، دانشکده مهندسی صنایع، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران
۲. دانشجوی کارشناسی‌ارشد، گروه فناوری اطلاعات، دانشکده مهندسی صنایع، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران ایران

تاریخ دریافت: ۱٤۰۳/۱۲/۲۰ تاریخ پذیرش: ۱٤۰٤/۰٦/۱۷

چکیده

با گسترش روزافزون خرید آنلاین و نقش پررنگ رسانه‌های اجتماعی در تصمیم‌گیری مصرف‌کنندگان، پیش‌بینی قصد خرید به یکی از مسائل کلیدی بازاریابی دیجیتال تبدیل شده است. پیش‌بینی دقیق این رفتار می‌تواند به بهبود هدف‌گیری تبلیغات، افزایش نرخ تبدیل مشتریان و بهینه‌سازی استراتژی‌های فروش کمک کند. هدف این پژوهش ارائه یک مدل ترکیبی مبتنی بر یادگیری ماشین برای پیش‌بینی قصد خرید کاربران با بهره‌گیری از بازخورد آنان نسبت به تبلیغات است. داده‌ها شامل ۱۰۰۰ رکورد و ۱۰ ویژگی مرتبط (از جمله سن، جنسیت، درآمد، مدت‌زمان حضور در سایت، میزان مصرف اینترنت روزانه، عنوان تبلیغ، محل و زمان تعامل و برچسب کلیک) بوده که از منابع عمومی (Kaggle) استخراج و پاک‌سازی شده‌اند. پس از پیش‌پردازش و انتخاب ویژگی‌ها، از روش پشته‌سازی (stacking) استفاده شد که در آن خروجی مدل‌های پایه (رگرسیون لجستیک و ماشین بردار پشتیبان) ترکیب و به‌عنوان پیش‌بینی نهایی به‌کار رفت. ارزیابی مدل با معیارهای صحت، دقت، یادآوری و F۱ نشان داد که مدل پیشنهادی عملکرد مناسبی دارد (صحت ۰.۹٦ و دقت ۰.۹۸). یافته‌ها بیانگر آن است که ترکیب ویژگی‌های دموگرافیک و رفتار کاربر می‌تواند به بهبود پیش‌بینی قصد خرید کمک نماید و برای هدف‌گیری تبلیغات در پلتفرم‌های تجارت الکترونیک مفید واقع شود.

کلیدواژه‌ها: قصد خرید، رفتار مشتری، رسانه‌های اجتماعی، بازاریابی

---

* نویسنده مسئول مقاله E-mail: hosseini@kntu.ac.ir

۱- مقدمه و بیان مسئله

تحقیقات همواره بر تجزیه و تحلیل رفتار مشتریان در فضاهای آنلاین تأکید می‌کند. اخیرا با رشد اینترنت و تبدیل‌شدن رسانه‌های اجتماعی به یکی از مهم‌ترین نقاط تماس شرکت‌ها با مشتریان اهمیت این قضیه بیشتر نیز می‌شود[۱]. پژوهش‌ها نشان می‌دهند که ارائه توصیه‌های یکسان برای همه مصرف‌کنندگان بالقوه یک استراتژی ناکارآمد است [۱]. طبق تحقیقات تعداد جلسات آنلاین که با خرید کاربر به پایان می‌رسد در مقایسه با تعداد کل جلسات آنلاین بسیار کم است و اکثریت بازدیدهای آنلاین بدون خرید رها می‌شوند [۲]. بنابرین تلاش‌های بازاریابی باید هم‌راستا با نیازها و ترجیحات مشتریان بالقوه صورت بگیرد تا مؤثر واقع شود [۳]. با پیش‌بینی احتمال قصد خرید هر مشتری، کسب‌وکارها می‌توانند استراتژی‌های بازاریابی خود را به طور مؤثرتری تنظیم کنند؛ توصیه‌های شخصی‌سازی شده ارائه دهند و در نهایت نرخ تبدیل مشتری و فروش را افزایش دهند [٤]. با پیش‌بینی دقیق خرید هر یک از مشتریان بازاریابی هدفمندتر صورت می‌گیرد؛ فروش افزایش می‌یابد؛ مدیریت موجودی کالا بهبود می‌یابد؛ هزینه‌ها کاهش و رضایت مشتریان نیز افزایش می‌یابد [٥]. برای پیش‌بینی رفتار خرید کاربران، لازم است عوامل مؤثر (از جمله ویژگی‌های دموگرافیک و الگوهای رفتاری در مواجهه با تبلیغات) شناسایی و مدل‌سازی شوند. این پژوهش بر توسعه چارچوبی ترکیبی برای پیش‌بینی احتمال خرید کاربران مبتنی بر ترکیب داده‌های رفتاری و جمعیت‌شناختی تمرکز دارد. تحقیقات اخیر در تلاش‌اند تا با استفاده از رفتاری که مشتری در فضاهای آنلاین دارد مثل جریان کلیک یا مدت زمان حضور در صفحات سایت، قصد خرید مشتریان را پیش‌بینی کند [۲]. به‌طور کلی، پیش‌بینی قصد خرید به‌عنوان یک ابزار ارزشمند برای درک رفتار مصرف‌کننده، بهینه‌سازی استراتژی‌های بازاریابی و افزایش درآمدزایی عمل می‌کند. با شناخت عوامل مؤثر بر قصد خرید و به کارگیری مدل‌های پیش‌بینی دقیق، شرکت‌ها می‌توانند تلاش‌های خود را با خواسته‌ها و ترجیحات مصرف‌کننده هماهنگ کنند. با توجه به مباحث مطرح‌شده، ضرورت توسعه مدل‌های دقیق‌تر برای پیش‌بینی قصد خرید بیش از پیش آشکار می‌شود. مرور مطالعات پیشین نشان می‌دهد که بیشتر پژوهش‌ها یا بر داده‌های نشست وب‌سایت‌ها (مانند جریان کلیک و مدت زمان حضور در صفحات) تمرکز داشته‌اند یا بر تحلیل محتوای

[۱] www.forbes.com

تولیدشده توسط کاربر (مانند نظرات و امتیازدهی). هرچند این رویکردها بینش ارزشمندی ایجاد کرده‌اند، اما غالبا از ترکیب ویژگی‌های دموگرافیک کاربران (مانند سن، جنسیت و سطح درآمد) با داده‌های رفتاری مرتبط با تعامل آن‌ها با تبلیغات (از جمله کلیک روی تبلیغ و مدت‌زمان حضور) غفلت شده است. به‌ویژه نقش عواملی چون سن، جنسیت و سطح درآمد در کنار شاخص‌های رفتاری در بسیاری از تحقیقات نادیده گرفته شده است. این خلأ پژوهشی ضرورت بررسی مدلی ترکیبی را آشکار می‌سازد که بتواند رفتار کاربر در مواجهه با تبلیغات را همراه با ویژگی‌های جمعیت‌شناختی او برای پیش‌بینی قصد خرید به کار گیرد.

پژوهش حاضر با هدف پر کردن این شکاف، یک مدل ترکیبی مبتنی بر الگوریتم‌های یادگیری ماشین ارائه می‌دهد که در آن داده‌های رفتاری مرتبط با تبلیغات (مانند کلیک و مدت‌زمان تعامل) در کنار ویژگی‌های جمعیت‌شناختی کاربران برای پیش‌بینی قصد خرید به‌کار گرفته می‌شود. نوآوری اصلی این پژوهش در ادغام داده‌های رفتاری مرتبط با تبلیغات و ویژگی‌های جمعیت‌شناختی در قالب یک چارچوب ترکیبی مبتنی بر روش پشته‌سازی است که امکان پیش‌بینی دقیق‌تر قصد خرید را فراهم می‌سازد. این پژوهش با هدف پاسخ به پرسش‌های زیر طراحی شده است:

- چگونه می‌توان با استفاده از ترکیب ویژگی‌های رفتاری (مانند کلیک و مدت‌زمان تعامل) و دموگرافیک کاربران، قصد خرید را پیش‌بینی کرد؟
- آیا به‌کارگیری یک مدل ترکیبی مبتنی بر روش پشته‌سازی می‌تواند عملکرد بهتری نسبت به مدل‌های منفرد در پیش‌بینی قصد خرید ارائه دهد؟ آیا یک مدل پیش‌بینی می‌تواند از بازخورد کاربران به تبلیغات، قصد خرید آن‌ها را پیش‌بینی کند؟

ساختار پژوهش حاضر به این صورت است: در بخش دوم مفاهیم کلی و مرور ادبیات مرتبط ارائه می‌شود. بخش سوم به تشریح مراحل و روش‌های تحقیق اختصاص دارد و در بخش چهارم پیاده‌سازی این مراحل انجام می‌گیرد. در بخش پنجم نیز جمع‌بندی، نتیجه‌گیری و پیشنهادهایی برای تحقیقات آتی مطرح خواهد شد.

۲- مروری بر ادبیات موضوع

به‌منظور انسجام بیشتر، مرور ادبیات این پژوهش در سه حوزه اصلی بررسی شده است:

۱- ویژگی‌های رفتاری و دموگرافیک مؤثر بر قصد خرید

۲- نقش رسانه‌های اجتماعی و تبلیغات آنلاین در شکل‌گیری قصد خرید

۳- مدل‌ها و الگوریتم‌های مبتنی بر یادگیری ماشین در پیش‌بینی قصد خرید

قصد خرید به احتمال یا تمایل مصرف‌کنندگان برای خرید محصول یا خدمات گفته می‌شود. این موضوع نقش پررنگی در استراتژی‌های بازاریابی کسب‌وکارها ایفا می کند؛ زیرا به آن‌ها کمک می‌کند تا علاقه مصرف‌کنندگان را بسنجند و تلاش‌های فروش خود را بهینه کنند [٤]. برای پیش‌بینی قصد خرید، محققان و بازاریابان از تکنیک‌ها و روش‌های علمی مختلفی استفاده می‌کنند؛ نظرسنجی و مصاحبه، تکنیک‌های مدل‌سازی آماری پیشرفته، مانند تحلیل رگرسیون، مدل‌سازی معادلات ساختاری و الگوریتم‌های یادگیری ماشین [٦, ٧]. قصد خرید آنلاین به‌عنوان میزان تمایل مصرف‌کننده به خرید یک محصول از طریق فروشگاه‌های آنلاین تعریف می‌شود [٨]. از نظر وو و همکاران[1] (۲۰۱۱)، در زمینه تبلیغات، قصد خرید را به فرآیندی اطلاق می‌کنند که در آن مصرف‌کنندگان به دلیل تبلیغات ارائه شده از طریق رسانه‌های اجتماعی، برنامه‌ای را برای خرید یک محصول یا خدمت در نظر می‌گیرند [٩]. تبلیغات رسانه‌های اجتماعی شکلی از تبلیغات اینترنتی هستند؛ اما در مقایسه با تبلیغات رسانه‌های جمعی سنتی و سایر تبلیغات آنلاین بسیار تعاملی‌تر هستند [١٠]. شرکت‌ها می‌توانند ارتباط آموزنده‌تر و تعاملی (دوطرفه) بیشتری با مشتریان خود در این فضا داشته باشند و کاربران به دلیل تعامل بیشتر در فضای رسانه‌های اجتماعی (مانند لایک کردن، اشتراک‌گذاری و نظر دادن) ادراک و تجربیات متفاوتی خواهند داشت و تبلیغات در این فضا هدفمندتر صورت خواهد گرفت. تلاش‌های بازاریابی باید هم‌راستا با نیازها و ترجیحات مشتریان بالقوه صورت بگیرد تا مؤثر واقع شود. پیش‌بینی قصد خرید راه‌حل این مسئله است. پیش‌بینی قصد خرید مشتری موجب می‌شود تا روش‌های بازاریابی یک کسب‌وکار هدفمندتر، مدیریت موجودی‌اش بهینه‌تر و در نتیجه هزینه‌هایش کمتر شود [٥]. با پیش‌بینی احتمال قصد خرید هر مشتری، کسب‌وکارها می‌توانند استراتژی‌های بازاریابی خود را به طور مؤثرتری تنظیم کنند؛ توصیه‌های شخصی‌سازی شده ارائه دهند و در نهایت نرخ تبدیل مشتری و فروش را افزایش دهند [٤]. در این بخش، مروری بر کوشش محققان قبلی در راستای پیش‌بینی رفتار خرید مشتریان می‌شود.

[1] Wu et al.

۲-۱- ویژگی‌های رفتاری و دموگرافیک

لامهاری و همکاران[۱] (۲۰۲۲)، ساختاری جهت حفظ، جذب و تبدیل مشتری ارائه دادند [٦]. آنان مدل خود را مدیریت ارتباط با مشتری اجتماعی تحلیلی نام‌گذاری کردند که به جمع‌آوری صدای مشتریان[۲] از پلتفرم‌های رسانه‌های اجتماعی جهت بهبود حفظ و جذب آن‌ها کمک می‌کند. طبق مدل لامهاری از تحلیل احساسات روی کامنت‌های کاربران در رسانه‌های اجتماعی، نیازهای آن‌ها دسته‌بندی می‌شود و در مرحله بعد مشتریان در سه کلاس مجزا "ناراضی"، "راضی" و "بالقوه" طبقه‌بندی می‌شوند. پس از این مرحله سیاست‌های جذب و حفظ مشتری متناسب با هر یک از دسته‌ها صورت خواهد گرفت. وانگ و همکاران[۳] (۲۰۲۰)، روشی برای الگوهای پیش سفارش مشتری ابداع کردند [۱۱]. این روش مبتنی بر بخش‌بندی بازار[٤] و مدل RFM است. طبق راه‌حل پیشنهادی وانگ، مشتریان طبق شباهت‌هایشان در دسته‌های متفاوتی تقسیم‌بندی شده و سپس الگوی خرید گذشته‌شان بررسی می‌شود و به کمک الگوریتم‌های پیش‌بینی، رفتار خرید آینده هر گروه پیش‌بینی می‌شود. چو و همکاران[٥] (۲۰۱۹) با استفاده از داده‌های جریان کلیک کاربران، ساختاری مبتنی بر یادگیری ماشین را اجرا کردند که خریدار بودن یا نبودن یک کاربر را مشخص می‌کند [۱۲]. همچنین پیش‌بینی می‌کند که کاربر به یک محصول علاقه دارد یا خیر و امکان خرید آن در آینده وجود دارد یا خیر؟ روش پیشنهادی چو، مبنی بر شبکه عصبی بازگشتی[٦] و الگوریتم نزدیک‌ترین همسایه[۷] است. باگ و همکاران[۸] (۲۰۱۹)، برای پیش‌بینی قصد خرید کالاهای بادوام، روشی بر پایه سیستم‌های توصیه‌گر پیشنهاد دادند [۱۳]. طبق مدل آن‌ها با تحلیل احساسات نظرات مشتریان و کاویدن شبکه‌های اجتماعی، محصولات مورد علاقه هر مشتری شناسایی شده و توصیه مؤثر و مرتبط با آن به آن‌ها ارائه می‌شود.

با وجود مطالعات متعدد در زمینه ویژگی‌های رفتاری و دموگرافیک، بسیاری از پژوهش‌ها بر یک نوع داده یا یک رویکرد خاص متمرکز بوده‌اند و تعامل هم‌زمان ویژگی‌های دموگرافیک و رفتار کاربران کمتر مورد بررسی قرار گرفته است؛ بنابراین نیاز به چارچوب‌های ترکیبی محسوس است.

---

[۱] Lamrhari et al.
[۲] Voice of the customer (VoC)
[۳] Wang et al.
[٤] Market Segmentation
[٥] Chu et al.
[٦] Recurrent Neural Network (RNN)
[۷] K-Nearest Neighbors (KNN)
[۸] Bag et al.

۲-۲- نقش رسانه‌های اجتماعی و تبلیغات آنلاین

الالوان باتکیه‌بر تئوری یکپارچه پذیرش و استفاده از فناوری[۱]، عوامل کلیدی که بر قصد مشتریان در خرید محصولات ارائه شده در تبلیغات رسانه‌های اجتماعی تأثیرگذار است را برشمرده است [۱۶]. این موارد عبارت‌اند از: انتظار عملکرد، انگیزه لذت‌گرا، عادت، تعامل، اطلاع‌دهندگی و ارتباط درک شده. طبق نتایج این پژوهش، درصورتی‌که محتوای تبلیغات رسانه‌های اجتماعی از نظر مشتریان غنی از اطلاعات مرتبط با محصول باشد و تبلیغات متناسب با ترجیحات و علایق مشتری و به شیوه‌ای نوآورانه و خلاقانه صورت گیرد؛ می‌تواند مشتریان را در خرید محصولات به‌شدت ترغیب کند. آن‌ها همچنین دریافتند که عادت تأثیری در قصد خرید ندارد و تعامل و اطلاع‌دهندگی به ترتیب مهم‌ترین عوامل پیش‌بینی‌کننده قصد خرید هستند. نصیر و همکاران[۲] (۲۰۲۱)، پژوهشی با هدف شناسایی و دسته‌بندی بخش‌های مختلف مصرف‌کننده بر اساس ویژگی‌های تبلیغات رسانه‌های اجتماعی و بر پایه مطالعه الالوان انجام داده‌اند [۷]. آن‌ها در این پژوهش مصرف‌کنندگان را بر اساس تصوراتشان در مورد تبلیغات رسانه‌های اجتماعی به سه دسته "تأثیرپذیر"[۳]، "بی‌طرف"[۴] و "تأثیرناپذیر"[۵]. تقسیم‌بندی می‌کنند و سپس مشخص می‌کنند که قصد خرید هر دسته از مصرف‌کنندگان تحت‌تأثیر چه عواملی است. به طور مثال مصرف‌کنندگان "تأثیرپذیر" در مقایسه با مصرف‌کنندگان "بی‌طرف" و "تأثیرناپذیر"، تمایل به خرید انگیزشی بیشتری دارند و تمایل به تعامل بیشتری با صفحات رسانه‌های اجتماعی دارند و راحت‌تر تحت‌تأثیر تبلیغات قرار می‌گیرند. طبق نظر الالوان تعامل مهم‌ترین عامل پیش‌بینی‌کننده قصد خرید است. هم‌چنین مطابق با یافته‌های نوروزی و همکاران (۲۰۲۳)، استفاده از تحلیل احساسات برای پیش‌بینی قصد خرید کاربران در شبکه‌های اجتماعی می‌تواند دقت مدل‌های پیش‌بینی را بهبود دهد. نتایج پژوهش آنها در زمینه کاربران توییتر نشان داد که شاخص‌های رفتاری و محتوایی می‌توانند به‌طور معناداری بر قصد خرید اثرگذار باشند [۱۷]. برای اولین‌بار اما سریجش و همکاران[۶] (۲۰۲۰)، به طور تجربی اثبات کردند که لزوما تعامل تحت هر شرایطی، نتایج مثبت نخواهد داشت [۱۸]. طبق پژوهش آن‌ها، تعامل بسیار بالا در رسانه‌های اجتماعی بر اثربخشی تبلیغات

---

[۱] UTAUT۲
[۲] Nasir et al.
[۳] susceptible
[۴] dispassionate
[۵] impervious
[۶] Sreejesh et al.

تأثیر منفی می‌گذارد؛ زیرا فرصت پردازش مفهوم پیام و درک کافی را بسیار کمتر خواهد کرد. سریجش در پژوهش خود پیشنهاد کرد که اثر نامطلوب تعامل بالا را می‌توان با استفاده از واضح[۱] بودن پیام و آنتروپومرفیسم[۲] (تصور مفهوم و شخصیت انسانی برای چیزی) بهبود بخشید. دنگ و همکاران[۳] (۲۰۲۱)، پژوهشی مبنی بر اینکه سبک پست‌های شبکه‌های اجتماعی روی تعامل مصرف‌کنندگان تأثیر دارد یا خیر انجام دادند [۱۹]. آن‌ها با متن‌کاوی صفحات فیس‌بوک ۱۰٤ برند مطرح به این نکته اشاره کردند که جذابیت عاطفی، پیچیدگی زبانی و رسمی بودن یا نبودن پست‌ها در تعداد لایک، کامنت و اشتراک‌گذاری‌ها بسیار مؤثر است. نوع تصویر و لحن گوینده تبلیغات نیز عوامل بسیار مهمی هستند که چانگ و همکاران[٤] (۲۰۱۹) در پژوهش خود به آن اشاره کردند [۲۰]. نتایج پژوهش آنان نشان می‌دهد که استفاده از لحن متناسب با تصویر هر برند بسیار مناسب مهم است. در صورت که تصویر برند گرم و صمیمی باشد؛ استفاده از روایت اول‌شخص موجب جلب‌توجه بیشتر مشتریان برند و افزایش انگیزه‌های تعلق اجتماعی آن‌ها می‌شود؛ اما درصورتی‌که تصویر برند تصویری لوکس و شایسته باشد باید از لحن سوم‌شخص استفاده کرد. دونگ و همکاران[۵] (۲۰۱۹)، در پژوهشی به تأثیر مهم مصرف‌کنندگان بر میزان متقاعد کنندگی یکدیگر اشاره کردند [۲۱]. ازآنجاکه رسانه‌های اجتماعی فضای بسیار غنی برای تبادل اطلاعات C۲C است؛ شرکت‌ها باید از این فرصت برای انتشار شهرت مثبت و فرصت تبلیغات الکترونیکی دهان‌به‌دهان استفاده کنند. در پژوهشی از پاتل و همکاران[٦] (۲۰۲۰)، تأثیر کیفیت رابط کاربری اپلیکیشن‌های فروش، بررسی شد [۲۲]. کیفیت رابط کاربری در سه بعد که عبارت‌اند از میزان اطلاعات (کلی/در مورد محصولات)، چیدمان آیتم‌ها و جذابیت بصری سنجیده شد. نتایج این مطالعه نشان می‌دهد که کیفیت رابط کاربری برنامه‌های خرید تأثیر مستقیمی بر لذت درک شده مشتری دارد و همچنین تأثیر مثبتی بر قصد خرید نیز دارد؛ اما لزوما رابطه بین قصد خرید و کیفیت رابط کاربری مستقیم نیست. همچنین در ادامه پژوهش مشخص شد که مهم‌ترین بعد کیفیت رابط کاربری اطلاعات ارائه شده است. مک کلور و همکاران[۷] (۲۰۲۰) نیز دریافتند که هم آشنایی با برند و هم کیفیت اطلاعات

---

[۱] vividness
[۲] Anthropomorphism
[۳] Deng et al.
[٤] Chang et al.
[۵] Duong et al.
[٦] Patel et al.
[۷] McClure et al.

محتوای رسانه‌های اجتماعی آن به طور قابل‌توجهی بر مشارکت مصرف‌کنندگان مؤثر است [۲۳]. آنان همچنین کیفیت اطلاعات محتوا را بیشتر از آشنایی مشتری با نام برند در قصد خرید مؤثر دانستند و این مهم فرصت بسیار مناسبی برای نام‌های تجاری ناآشناست تا با ارائه اطلاعات باکیفیت و مناسب در صفحات اجتماعی‌شان، مشارکت مخاطبین خود را برانگیزند. بازی و همکاران[۱] (۲۰۲۰)، انگیزه مشتریان برای تعامل با برندها را در رسانه‌های اجتماعی بررسی کردند و مهم‌ترین عامل، ارتباط محتوای درک شده بود [۲٤]. طبق نظر آنان محتوای رسانه باید نیازهای اطلاعاتی مشتریان را برآورده سازد و از کیفیت زیباشناختی و جذابیت بصری بالایی نیز برخوردار باشد. همچنین آن‌ها در پژوهش خود از مزیت استفاده افراد مشهور در پست‌های برندهای مطرح سخن گفتند؛ زیرا مشتریان علاقه‌مند هستند که بدانند افراد مشهور چگونه از محصولات آنان استفاده می‌کنند.

در حوزه تبلیغات و رسانه‌های اجتماعی، شواهد قوی‌تری مبنی بر تأثیر محتوا، تعامل و کیفیت اطلاعات بر قصد خرید وجود دارد، اما بسیاری از مطالعات توصیفی‌اند، محدود به یک یا چند پلتفرم خاص یا فاقد تحلیل علیّت صریح؛ لذا نتایج نیازمند بررسی‌های تجربی و تعمیم‌پذیری بیشتر است.

۲-۳- مدل‌ها و الگوریتم‌های یادگیری ماشین

کبیر و همکاران[۲] (۲۰۱۹)، برای ایجاد یک مدل پیش‌بینی بهتر، عملکرد چند الگوریتم طبقه‌بندی نظارت‌شده و چند روش یادگیری جمعی[۳] را به بوته آزمایش گذاشتند [۱٤]. آن‌ها با استفاده از یک دیتاست ثابت که شامل داده‌های نشست کاربران بود؛ جنگل تصادفی را از سایر روش‌ها مناسب‌تر دیدند. کارتیک و همکاران[٤] (۲۰۲۱)، یک سیستم توصیه‌گر مبتنی بر منطق فازی را اجرا کردند [۱٥]. در روش آن‌ها، ابتدا کاربران هدف بر اساس تجزیه و تحلیل احساسات شناسایی می‌شدند؛ سپس میزان مشابهت هر دسته از کاربران با هم محاسبه گشته و در آخر هر گروه از کاربران بر اساس اطلاعات پروفایلی مشابه در دسته‌بندی مشابه قرار خواهد گرفتند و در نهایت با اعمال منطق فازی، سیستم لیستی از مرتبط‌ترین محصولات را به کاربران توصیه می‌کرد. اسمعیلی و همکاران[٥] (۲۰۲۱)، تابع امتیازدهی EPP را برای ارزیابی روش‌های

---

[۱] Bazi et al.
[۲] Kabir et al.
[۳] Ensemble learning
[٤] Karthik et al.
[٥] Esmeli et al.

پیش‌بینی قصد خرید مشتری ایجاد کردند [۲]. سیستم امتیازدهی آن‌ها، میزان پیش‌بینی صحیح سایر مدل‌های پیشین را مقایسه و اندازه‌گیری می‌کند. طبق گفته اسمعیلی و همکاران (۲۰۲۱) این روش در مواقعی کاربرد دارد که اطلاعات قبلی از کاربر نداریم و ناشناس است. طبق نتایج این پژوهش بهترین مدل برای پیش‌بینی قصد خرید درخت تصمیم است. الگاویاز و همکاران[۱] (۲۰۱۹)، الگوریتم تقویت تطبیقی را برای پیش‌بینی قصد خرید مشتری مناسب می‌دانند [٤]. آن‌ها مواردی را که می‌توان برای پیش‌بینی قصد خرید کاربر استفاده کرد را به دو دسته رفتارهای کاربر و جمعیت دسته‌بندی کردند.

در سال‌های اخیر، مطالعات متعددی به بررسی تأثیر عوامل مبتنی بر هوش مصنوعی و بازاریابی دیجیتال بر قصد خرید مصرف‌کنندگان پرداخته‌اند. برای مثال، اولاه و همکاران[۲] (۲۰۲۵) نشان دادند که شخصی‌سازی مبتنی بر هوش مصنوعی، دقت پیش‌بینی، و تعامل لحظه‌ای نقش معناداری در تقویت قصد خرید کاربران دیجیتال ایفا می‌کنند [۲۵]. به‌طور مشابه، السعدی و همکاران[۳] (۲۰۲۵) با استفاده از داده‌های جمع‌آوری‌شده از کاربران در عربستان، اثربخشی الگوریتم‌های یادگیری ماشین از جمله جنگل تصادفی و طبقه‌بندی‌کننده رای‌گیر را در پیش‌بینی قصد خرید کاربران دیجیتال با دقتی حدود ۹۰ درصد گزارش کرده‌اند [۲٦]. این یافته‌ها نشان می‌دهند که الگوریتم‌های طبقه‌بندی می‌توانند ابزاری مؤثر برای تحلیل رفتار خرید در محیط‌های دیجیتال باشند. همچنین در زمینه محصولات سبز، دنیس و همکاران[٤] (۲۰۲۵) بر اهمیت متغیرهایی نظیر کیفیت ادراک‌شده، ارزش ادراک‌شده و اعتماد سبز در شکل‌گیری نیت خرید تأکید کرده و نقش تقویتی بازاریابی سبز را نیز مورد تأیید قرار داده‌اند [۲۷]. مرور این مطالعات حاکی از آن است که ترکیب رویکردهای بازاریابی، فناوری‌های نوین، و تحلیل داده‌های رفتاری می‌تواند بینش‌های مؤثری برای درک و پیش‌بینی قصد خرید مصرف‌کنندگان ارائه دهد.

در حوزه روش‌شناسی و الگوریتم‌ها، مدل‌های ترکیبی و روش‌های جمعی غالباً عملکرد مناسبی نشان داده‌اند؛ با این حال در بسیاری از مطالعات گزارش مراحل پیش‌پردازش ناقص است،

---

[۱] Algawiaz et al.
[۲] Ullah et al.
[۳] Alsaadi et al.
[٤] Denis et al.

اعتبارسنجی محدود انجام شده و استفاده از مجموعه‌داده‌های کوچک یا خاص مانع از تعمیم‌پذیری نتایج می‌شود.

مرور ادبیات نشان می‌دهد که هرچند پژوهش‌های متعددی به بررسی قصد خرید پرداخته‌اند، اما بیشتر آن‌ها یا بر داده‌های نشست وب و رفتارهای کلیکی تمرکز داشته‌اند یا به تحلیل محتوای تولیدشده توسط کاربر پرداخته‌اند. اگرچه این رویکردها بینش‌های ارزشمندی فراهم کرده‌اند، اما غالبا با محدودیت همراه بوده‌اند و از ترکیب ابعاد مختلف، به‌ویژه بازخورد مستقیم کاربران نسبت به تبلیغات در کنار ویژگی‌های دموگرافیک، غفلت کرده‌اند. این ارزیابی انتقادی نشان می‌دهد که ادبیات موجود فاقد چارچوبی جامع و ترکیبی است که بتواند هم‌زمان داده‌های رفتاری، دموگرافیک و بازخورد تبلیغاتی را برای پیش‌بینی قصد خرید به کار گیرد. پژوهش حاضر با ارائه یک مدل ترکیبی مبتنی بر یادگیری ماشین، در پی پر کردن این شکاف و ارتقای دقت پیش‌بینی قصد خرید است.

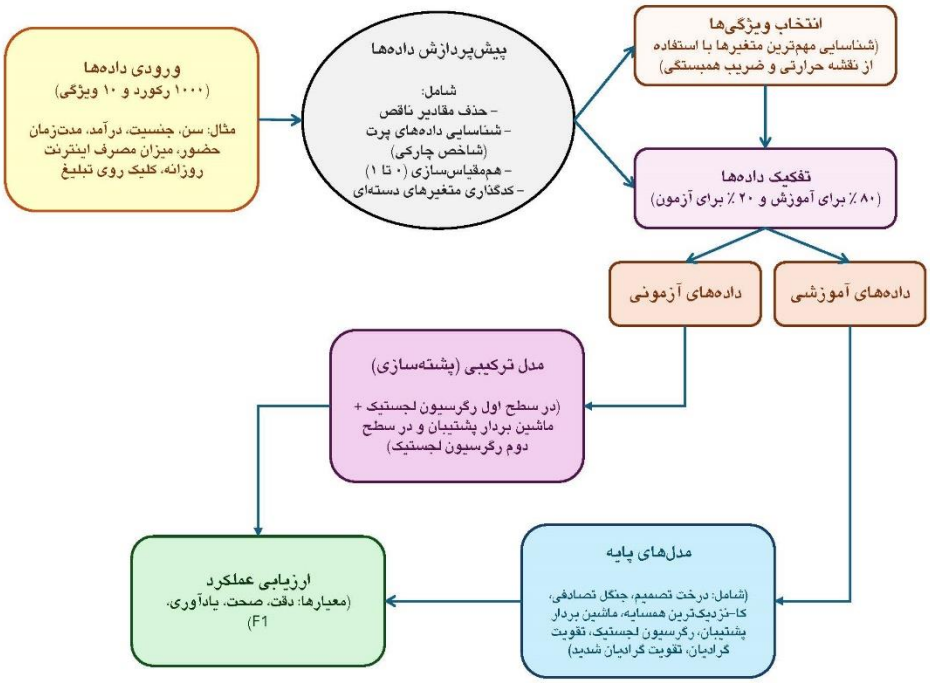
۳- روش تحقیق / چارچوب پژوهش

در این بخش، گام‌های گردآوری پژوهش تشریح می‌شوند. ابتدا به بررسی مجموعه‌داده مربوط به پژوهش و مرحله پیش‌پردازش می‌پردازیم. سپس مراحل انتخاب ویژگی و تفکیک داده انجام می‌شود و مدل پیش‌بینی پیشنهادی مورد بررسی قرار می‌گیرد. شکل ۱ مسیر چارچوب پژوهش را نشان می‌دهد. شکل چارچوب پیشنهادی پژوهش را نشان می‌دهد. این چارچوب شامل چهار مرحله اصلی است: ۱) پیش‌پردازش داده‌ها؛ ۲) آموزش مدل‌های پایه (رگرسیون لجستیک و ماشین بردار پشتیبان)؛ ۳) انتقال خروجی مدل‌های پایه به متامدل در سطح دوم پشته‌سازی؛ ٤) ارزیابی عملکرد مدل نهایی

این مراحل جریان داده‌ها از ورودی تا پیش‌بینی نهایی قصد خرید را به‌صورت یکپارچه نمایش می‌دهند. پس از ورود مجموعه‌داده متناسب با هدف پژوهش، مراحل پیش‌پردازش داده انجام می‌شود و داده‌ها به دو مجموعه آزمایشی و آموزشی تفکیک می‌گردند. برای بهبود دقت نتایج و کاهش بیش‌برازش[۱]، از روش یادگیری جمعی که ترکیبی از دو یا چند مدل پایه یادگیری ماشین است، استفاده کرده‌ایم.

---

[۱] Overfitting



شکل ۱: چارچوب پیشنهادی پژوهش

۳-۱- مجموعه‌داده

مجموعه‌داده مورد استفاده در این پژوهش، اطلاعات مرتبط با رفتار کاربران در مواجهه با تبلیغات یک سایت است. این دیتاست شامل ۱۰ ستون و ۱۰۰۰ سطر است. هر یک از سطرها اطلاعات منحصربه‌فرد یک کاربر را شامل می‌شود و هر ستون نوع متفاوتی از اطلاعات را از کاربر در اختیار می‌گذارد. این دیتاست شامل اطلاعات ۱۰۰۰ کاربر بوده و از منابع عمومی در دسترس (پلتفرم Kaggle) استخراج و پس از پاک‌سازی برای پژوهش حاضر مورد استفاده قرار گرفته است. این مجموعه‌داده از ده ستون تشکیل شده است که در جدول ۱ لیست شده‌اند.

جدول ۱: ستون های مجموعه‌داده مورد استفاده در این پژوهش

ستون‌ها	تعریف	ستون‌ها	تعریف
Daily Time Spent on Site	مدت زمانی که کاربر در سایت می‌گذراند.	City	شهر کاربر

ستون‌ها	تعریف	ستون‌ها	تعریف
Age	سن کاربر	Male	جنسیت کاربر
Area Income	میزان درآمد کاربر	Country	کشور کاربر
Daily Internet Usage	میزان مصرف اینترنت کاربر	Timestamp	زمانی که کاربر روی تبلیغ کلیک کرده است.
Ad Topic Line	عنوان تبلیغی که کاربر روی آن کلیک کرده است	Clicked on Ad	کلیک‌کردن یا نکردن روی تبلیغات

۳-۲- پیش‌پردازش مجموعه‌داده

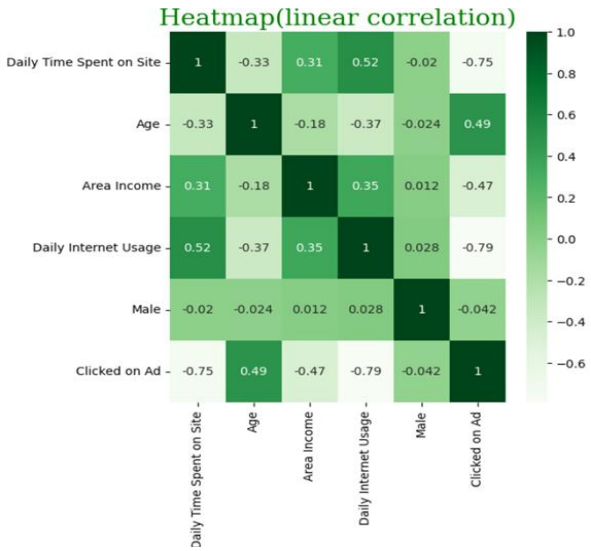
پیش‌پردازش داده‌ها یکی از مراحل ضروری در یادگیری ماشین است. در این پژوهش ابتدا رکوردهای دارای مقادیر ناقص بررسی و داده‌های پرت با استفاده از شاخص فاصله بین چارکی (IQR[۱]) شناسایی و مدیریت شدند. سپس برای هم‌مقیاس‌سازی متغیرهای عددی از روش مقیاس‌بندی Min–Max استفاده گردید تا تمامی مقادیر در بازه [۰,۱] قرار گیرند. علاوه بر این، متغیرهای دسته‌ای مانند جنسیت و عنوان تبلیغ نیز به‌صورت مناسب کدگذاری شدند. این فرایند موجب افزایش کیفیت و قابلیت اعتماد داده‌ها برای تحلیل و مدل‌سازی شد.

۳-۳- انتخاب ویژگی

در مرحله انتخاب ویژگی، از هیت‌مپ[۲] استفاده می‌کنیم که روشی مناسب برای اندازه‌گیری وابستگی بین ستون‌های مجموعه‌داده و شناسایی متغیرهای مرتبط با متغیر هدف است. این روش همچنین ارتباط متغیرهای پیش‌بینی را اندازه‌گیری کرده و در تشخیص کیفیت داده‌ها و حذف داده‌های غیرمرتبط[۳] مؤثر است. طبق هیت‌مپ شکل ۲، عنوان تبلیغات، مکان جغرافیایی کاربر و زمان ورود به سایت ارتباط کمی با سایر متغیرها دارند و بنابراین تأثیر زیادی بر خروجی الگوریتم‌های یادگیری ماشین نخواهند داشت. همچنین هیت‌مپ نشان می‌دهد که کلیک‌کردن یا نکردن روی تبلیغات، متغیر هدف ما است؛ زیرا بیشترین همبستگی را با سایر متغیرها دارد.

---

[۱] Interquartile Range
[۲] heatmap
[۳] anomaly detection



شکل ۲: هیت‌مپ پژوهش

جدول ۲: نتیجه مرحله انتخاب ویژگی

ویژگی‌های ورودی به مدل	"میزان زمان سپری شده در سایت"
	"میزان ترافیک اینترنت مصرفی کاربر در طول روز"
	"سن"
	"جنسیت"
	"درآمد"
متغیر هدف	"کلیک‌کردن یا نکردن روی تبلیغات"

این ویژگی‌ها بر اساس مطالعات پیشین و شواهد تجربی، از مهم‌ترین عوامل مؤثر بر رفتار خرید آنلاین کاربران هستند. برای مثال، زمان بیشتری که کاربر در سایت سپری می‌کند معمولاً با درگیری بیشتر ذهنی در فرآیند خرید همراه است، و درآمد نیز با توان خرید و تصمیم‌گیری اقتصادی رابطه مستقیم دارد. انتخاب این متغیرها مبتنی بر پیش‌فرضی است که رفتارهای دیجیتال کاربران بازتاب‌دهنده گرایش‌های واقعی آن‌ها به خرید است. نتیجه مرحله انتخاب ویژگی در جدول ۲ قابل مشاهده است. در این بخش به‌تفصیل به هر یک از این ویژگی‌ها و تاثیر آن‌ها در ورود به مدل و آموزش داده‌ها خواهیم پرداخت.

- **میزان زمان سپری شده در سایت**: این ویژگی میزان زمان بازدید هر کاربر از صفحات سایت را نشان می‌دهد. کاربران با هدف خرید معمولاً مدت زمان بیشتری را در سایت می‌گذرانند، چرا که به بررسی بیشتر محصولات و مطالعه نظرات دیگران می‌پردازند. در پژوهش‌های اخیر از این ویژگی برای پیش‌بینی قصد خرید استفاده شده است [۲, ٦]. در این پژوهش، غالب کاربران مدت زمانی بین ۵۱ تا ۸۲ دقیقه در سایت مذکور حضور داشته‌اند.
- **میزان ترافیک اینترنت مصرفی کاربر در طول روز**: این ویژگی میزان ترافیک اینترنت مصرفی کاربر در طول روز را مشخص می‌کند. عموماً افرادی که فعالیت کمتری در فضای مجازی دارند، تمایل کمتری به خرید آنلاین از خود نشان می‌دهند. طبق پژوهش نصیر و همکاران (۲۰۲۱)، ٤۲ درصد از مصرف‌کنندگان "مشتاق" روزانه بین ۳ تا ٤ ساعت از رسانه‌های اجتماعی استفاده می‌کنند، درحالی‌که حدود ۷۵ درصد از گروه "غیر مشتاق" کمتر از دو ساعت در روز از رسانه‌های اجتماعی استفاده می‌کنند [۷]. در تحقیق ما، میزان اینترنتی مصرفی روزانه کاربران سایت بین ۱۳۸ تا ۲۱۹ مگابایت است.
- **سن**: سن کاربر نیز بر رفتار خرید او تأثیرگذار است. در پژوهش نصیر و همکاران (۲۰۲۱)، ٪۳۹ خریداران در بازه سنی ۲۵ تا ۳۰ سال و ٤۳٪ غیر خریداران در بازه سنی ۳۱ تا ٤۰ سال بودند [۷]. در پژوهش ما، غالب کاربران در بازه سنی ۲۸ تا ٤۳ سال قرار دارند.
- **جنسیت**: سن از عوامل تأثیرگذار بر قصد خرید است و این ویژگی با متغیر دیگری چون میزان زمانی که کاربران در سایت می‌گذرانند؛ مرتبط است. نتایج نشان می‌دهد که مردان به طور متوسط زمان بیشتری نسبت به زنان در سایت می‌گذرانند.
- **درآمد**: درآمد یک مشتری قدرت خرید او را معین می‌کند و به پیش‌بینی قصد خرید کمک می‌کند. طبق پژوهش نصیر و همکاران (۲۰۲۱)، اکثریت کاربران در گروه "تأثیرناپذیر" کم‌درآمد (٤۲٪) و از گروه درآمد متوسط (۳۱٪) هستند. درحالی‌که پاسخ‌دهندگان در خوشه "تأثیرگذار" بیشتر از گروه‌های درآمدی متوسط روبه‌بالا یا بالا (۳۷٪) بودند [۷]. در این پژوهش میزان درآمد متوسط کاربران سایت، در بازه ٤۷۰۰۰ دلار تا ٦۷۰۰۰ دلار است.
- **کلیک‌کردن/نکردن روی تبلیغات**: کلیک‌کردن و جریان کلیک یکی از مهم‌ترین شاخص‌ها در سنجش خریدار بودن کاربر است. عدم کلیک بر روی تبلیغات نشان می‌دهد که کاربر رغبتی به خرید محصول ندارد و حتی به دیدن تبلیغات هم راغب نیست. در پژوهش چو و همکاران

(۲۰۱۹) و همچنین اسمعیلی و همکاران (۲۰۲۱) از این معیار استفاده شده است [۲, ۱۲]. دیتاست این پژوهش شامل ۱۰۰۰ کاربر است که نیمی از آن‌ها بر بنر تبلیغات کلیک کرده و نیمی دیگر کلیک نکرده‌اند. پژوهش ما نشان می‌دهد که هرچه کاربران مدت زمان بیشتری در سایت بگذرانند، احتمال کلیک آن‌ها روی تبلیغات کمتر می‌شود. همچنین میزان اینترنت مصرفی روزانه کاربرانی که روی تبلیغات کلیک می‌کنند نسبت به آن‌هایی که کلیک نمی‌کنند، کمتر است. این دو نکته جالب است زیرا برخلاف تصور، انتظار می‌رود که کاربران فعال‌تر بیشتر کلیک کنند، ولی یافته‌ها نشان می‌دهد که این‌طور نیست. همچنین، اکثریت کاربرانی که روی تبلیغات کلیک می‌کنند، در رنج درآمدی متوسط هستند و کاربرانی که درآمد بالاتری دارند، تمایل کمتری به کلیک دارند.

۳-۴- الگوریتم‌های یادگیری ماشین

- **درخت تصمیم**: درخت تصمیم یک الگوریتم یادگیری نظارت شده محبوب است که جهت طبقه‌بندی داده‌ها و رگرسیون استفاده می‌شود. هر گره، یک ویژگی و هر شاخه، یک تصمیم‌گیری و هر گرهٔ برگ نشان‌دهنده نتیجه یا برچسب کلاس‌ها هستند. این الگوریتم برای طبقه‌بندی و رگرسیون به‌کار می‌رود [۲۸]. اسمعیلی و همکاران (۲۰۲۱) آن را مناسب‌ترین روش برای پیش‌بینی قصد خرید یافته‌اند [۲]. ما نیز در این پژوهش از آن به‌عنوان طبقه‌بندی‌کننده داده‌های خود استفاده کرده‌ایم.
- **جنگل تصادفی**: جنگل تصادفی نوعی روش یادگیری جمعی است که با ایجاد مجموعه‌ای از درخت‌های تصمیم و ترکیب پیش‌بینی آن‌ها، عملکرد کلی را بهبود می‌بخشد. در هر درخت، انتخاب ویژگی‌ها به‌صورت تصادفی انجام می‌شود که موجب افزایش دقت و کاهش بیش‌برازش می‌گردد. در طول پیش‌بینی، هر درخت در جنگل به طور مستقل، یک پیش‌بینی تولید می‌کند و پیش‌بینی نهایی با رأی اکثریت یا میانگین‌گیری پیش‌بینی‌های تک درختان تعیین می‌شود [۲۹]. کبیر و همکاران (۲۰۱۹)، جنگل تصادفی را جهت پیش‌بینی قصد خرید مناسب دیدند و ما نیز با مجموعه‌داده خود عملکرد این مدل را بررسی می‌کنیم [۱٤].
- **کا-نزدیک‌ترین همسایه**: الگوریتم کا-نزدیک‌ترین همسایه یک روش یادگیری بانظارت و مبتنی بر نمونه است که برای طبقه‌بندی یا رگرسیون استفاده می‌شود. این الگوریتم بدون ساخت مدل از پیش، برای هر داده‌ی جدید، برچسب کلاس را بر اساس بیشترین تعداد رأی در

بین K نمونه نزدیک از داده‌های آموزشی تعیین می‌کند. معیار شباهت معمولاً فاصله اقلیدسی است. برای کارهای طبقه‌بندی، برچسب کلاس پیش‌بینی با اکثریت رأی، کا-نزدیک‌ترین همسایه تعیین می‌شود. برای کارهای رگرسیون، مقدار پیش‌بینی شده به‌عنوان میانگین مقادیر هدف کا-نزدیک‌ترین همسایگان محاسبه می‌شود [۳۰]. چو و همکاران (۲۰۱۹) در پژوهش خود از روش کا-نزدیک‌ترین همسایه برای پیش‌بینی خریدار بودن یا نبودن یک خریدار استفاده کردند و این روش عملکرد خوبی از خود نشان داده بود [۱۲]. ما نیز از این روش به‌عنوان طبقه‌بندی‌کننده استفاده می‌کنیم و در صورت عملکرد مطلوب آن را در مدل خود استفاده می‌کنیم.

- **ماشین بردار پشتیبان**:ماشین بردار پشتیبان یکی از الگوریتم‌های یادگیری نظارت‌شده است که به‌ویژه برای طبقه‌بندی داده‌های با ابعاد بالا کاربرد دارد. هدف آن یافتن یک ابرصفحه[۱] بهینه است که داده‌ها را با بیشترین فاصله ممکن از هم جدا کند و به تعمیم‌پذیری مدل کمک نماید. این الگوریتم در حوزه‌هایی مانند طبقه‌بندی متن، تصویر و بیوانفورماتیک بسیار مؤثر عمل می‌کند، اما در مجموعه‌داده‌های بزرگ‌تر، روش‌هایی مانند رگرسیون لجستیک یا شبکه‌های عصبی ممکن است کاراتر باشند. ویجایاراقوان و همکاران[۲] (۲۰۲۰)، جهت تجزیه و تحلیل احساسات نظرات محصولات آی‌پاد، یک مدل طبقه‌بندی مبتنی بر خوشه ارائه دادند [۳۱]. در مدل آن‌ها، جهت طبقه‌بندی بررسی محصول، از ماشین بردار پشتیبان استفاده می‌شود.
- **رگرسیون لجستیک**: این الگوریتم یادگیری نظارت‌شده برای طبقه‌بندی باینری کاربرد دارد و با استفاده از تابع سیگموئید، احتمال تعلق یک نمونه به یک کلاس خاص را بین ۰ تا ۱ پیش‌بینی می‌کند. برخلاف رگرسیون خطی، هدف آن پیش‌بینی احتمال است نه مقدار پیوسته. رگرسیون لجستیک به دلیل سادگی، در حوزه‌هایی مانند بازاریابی، امور مالی و سلامت به طور گسترده استفاده می‌شود و اغلب به‌عنوان مدل پایه در مقایسه با روش‌های پیچیده‌تر به کار می‌رود. ژینگ لو و همکاران[۳] (۲۰۱۸)، از لجستیک رگرسیون جهت کشف رابطه بین اطلاعات بازدید مشتریان از صفحات یک فروشگاه آنلاین محصولات مراقبتی پوست و تصمیمات خرید کاربران و تأثیر مثبت استفاده از لایو چت بر قصد خرید آن‌ها پرداختند [۳۲].

---

[۱] Hyperplane: برای یک مجموعه داده ۲ بعدی ساده، هایپرپلن یک خط است، در حالی که برای یک مجموعه داده سه بعدی، یک صفحه خواهد بود.

[۲] Vijayaragavan et al.

[۳] Lv et al.

- **تقویت گرادیان**: روش تقویت گرادیان، یک تکنیک یادگیری ماشینی جمعی قدرتمند است که برای کارهای رگرسیون و طبقه‌بندی استفاده می‌شود. داده‌ها را به روشی متوالی می‌سازد و نقاط قوت یادگیرندگان ضعیف مثل درخت‌های تصمیم‌گیری را ترکیب می‌کند تا یک مدل بسیار دقیق ایجاد کند. ایده کلیدی این روش است که هر مدل جدید به‌گونه‌ای ساخته شود که با تمرکز بیشتر بر روی نقاط داده‌ای که ضعیف پیش‌بینی شده بودند؛ خطاهای مدل‌های قبلی را تصحیح کند.
- **تقویت گرادیان شدید**[۱]: این الگوریتم نسخه‌ای بهینه‌شده از تقویت گرادیان است که با سرعت بالا و دقت زیاد، به‌ویژه در داده‌های ساختاریافته، عملکرد چشمگیری دارد. مانند روش اصلی، با افزودن تدریجی درختان تصمیم و اصلاح خطاهای قبلی، یک مدل قوی ایجاد می‌کند. پشتیبانی از محاسبات موازی، آن را در تحلیل‌های بازاریابی، مالی و سلامت به روشی مؤثر تبدیل کرده است. شهبازی و همکاران[۲] (۲۰۲۰)، با استفاده از الگوی یادگیری ماشین مبتنی بر تقویت گرادیان شدید، دقت سیستم توصیه‌گر یک فروشگاه آنلاین را بهبود دادند [۳۳]. این سیستم بر اساس الگوی کلیک کاربران، محصولات را به خریدارن پیشنهاد می‌کند. ژینگفن و همکاران[۳] (۲۰۱۸)، رگرسیون و تقویت گرادیان شدید را برای پیش‌بینی رفتار خرید کاربران در یک وب‌سایت تجارت الکترونیکی ترکیب می‌کنند [۳٤]. تحقیق آن‌ها نشان می‌دهد که ترکیب رگرسیون لجستیک و تقویت گرادیان شدید امکان پذیر بوده و شاخص ارزیابی مدل بهتر از سایر روش‌هایی است که به تنهایی استفاده می‌شود.

۴- یافته‌های پژوهش

۴-۱- انتخاب بهترین روش پیش‌بینی(مدل پایه)

در بخش قبل گام‌هایی که برای گردآوری این تحقیق انجام شد؛ عنوان شد و سپس چارچوب پیشنهادی تحقیق بررسی گشت. در این بخش به چگونگی پیاده‌سازی گام‌هایی این پژوهش خواهیم پرداخت. ارائه راه‌حل برای مسئله این پژوهش، نیاز به روش‌های یادگیری ماشین دارد؛ لذا برای پیاده‌سازی از برنامه‌نویسی به زبان پایتون[٤] استفاده می‌شود. پایتون به دلیل وجود

---

[۱] Extreme Gradient Boosting(XGBOOSTING)
[۲] Shahbazi et al.
[۳] XingFen et al.
[٤] Python

کتابخانه‌های کارآمد در حوزه یادگیری ماشین، بهترین انتخاب برای کار با کلان‌داده‌ها[1]ست. پس از ورود داده‌های پژوهش و انتخاب ویژگی، نوبت به تفکیک مجموعه‌داده می‌رسد. در این مرحله مجموعه‌داده به دو دسته داده‌های آموزش و داده‌های آزمایش تقسیم می‌شوند. داده‌های آموزش برای آموزش مدل استفاده می‌شوند. این داده‌ها بخش قابل‌توجهی از مجموعه‌داده را شامل می‌شوند و الگوها، روابط و روندهای بین داده‌ها را فرا می‌گیرند. از مجموعه آزمایش نیز برای ارزیابی و سنجش عملکرد مدل آموزش‌دیده استفاده می‌شود. در این پژوهش تفکیک مجموعه‌داده را با نسبت ۲۰/۸۰ انجام شد. ۸۰ درصد داده‌ها، داده‌های آموزشی و ۲۰ درصد آن‌ها جزء داده‌های آزمایشی بودند. مدل با استفاده از داده‌های آموزشی که حاوی ویژگی‌های منتخب هستند آموزش داده می‌شود (بخش ۳-۳) و سپس عملکرد مدل با استفاده از داده‌های آزمایش ارزیابی می‌شود.

در این پژوهش از روش یادگیری جمعی به‌منظور بهبود دقت نتایج و کاهش بیش‌برازش استفاده می‌کنیم. ما از تکنیک پشته‌سازی[2] به‌عنوان متامدل پژوهش استفاده خواهیم کرد و از روش‌هایی چون الگوریتم جنگل تصادفی، درخت تصمیم، کا-نزدیک‌ترین همسایه، ماشین بردار پشتیبان، رگرسیون لجستیک و تقویت گرادیان به‌عنوان مدل‌های پایه‌ای استفاده خواهیم کرد.

برای آموزش و ارزیابی مدل‌ها، از شش الگوریتم یادگیری ماشین شامل: درخت تصمیم، جنگل تصادفی، کا-نزدیک‌ترین همسایه، ماشین بردار پشتیبان، رگرسیون لجستیک، تقویت گرادیان و تقویت گرادیان شدید استفاده شد. پاسخ هر یک از الگوریتم‌ها به مجموعه‌داده و ویژگی‌های انتخاب‌شده از طریق معیارهای ارزیابی (صحت، دقت، یادآوری، F۱ و پشتیبانی) و ماتریس‌های سردرگمی تحلیل گردید. نتایج این تحلیل‌ها در جدول ۳ و شکل ۳ آورده شده است.

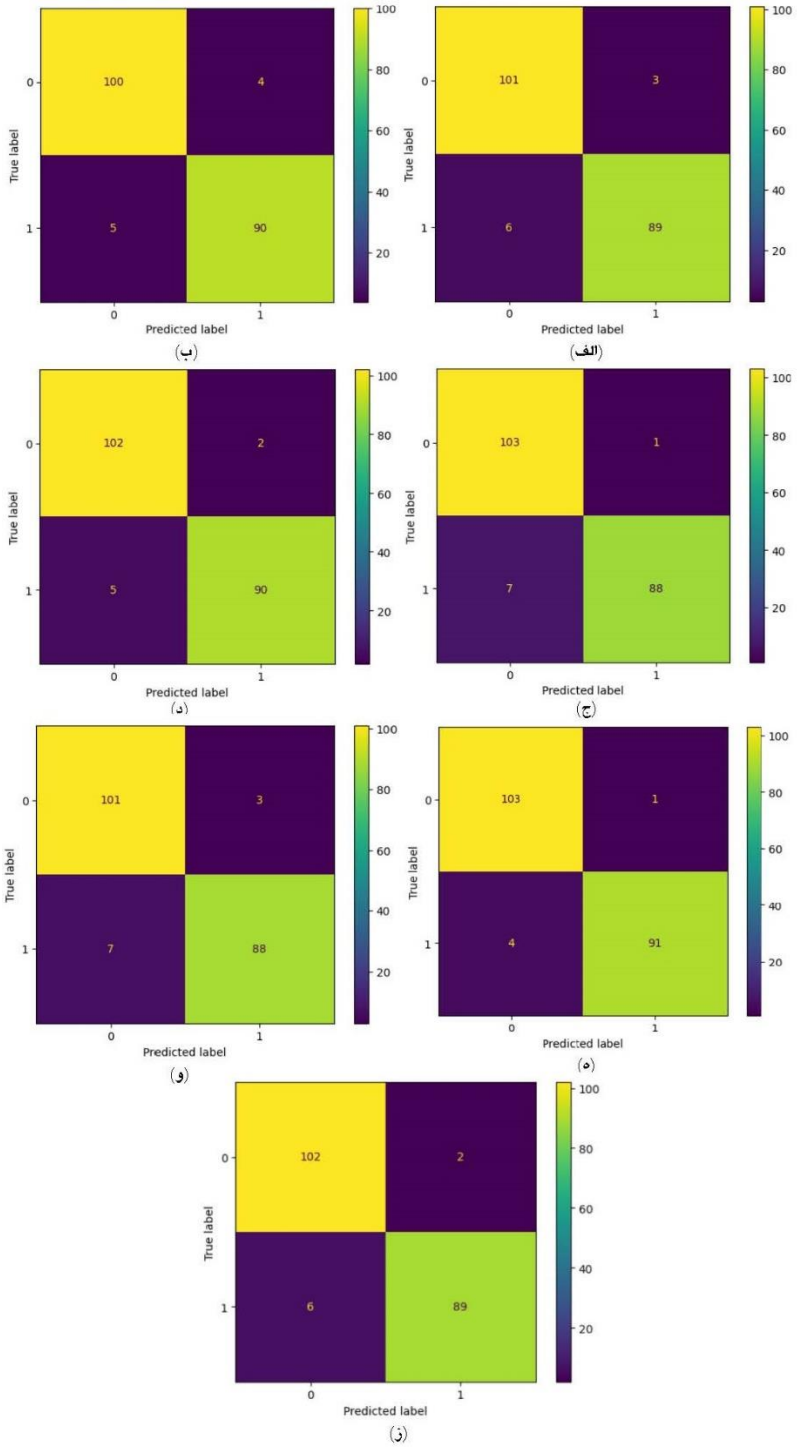
جدول ۳: نتایج ارزیابی طبقه‌بندی الگوریتم‌های یادگیری ماشین

الگوریتم یادگیری ماشین	برچسب	صحت	دقت	یادآوری	F۱	پشتیبانی
درخت تصمیم	برچسب ۱	۰.۹۵	۰.۹۷	۰.۹٤	۰.۹٥	۹٥
	برچسب ۰	۰.۹٥	۰.۹٤	۰.۹۷	۰.۹٦	۱۰٤
جنگل تصادفی	برچسب ۱	۰.۹٥	۰.۹٦	۰.۹٦	۰.۹٥	۹٥
	برچسب ۰	۰.۹٥	۰.۹٥	۰.۹٥	۰.۹٦	۱۰٤

[1] Big Data
[2] Stacking

الگوریتم یادگیری ماشین	برچسب	صحت	دقت	یادآوری	F۱	پشتیبانی
کا-نزدیک‌ترین همسایه	برچسب ۱	۰.۹۶	۰.۹۹	۰.۹۳	۰.۹۶	۹۵
	برچسب ۰	۰.۹۶	۰.۹٤	۰.۹۹	۰.۹۶	۱۰٤
ماشین بردار پشتیبان	برچسب ۱	۰.۹۶	۰.۹۸	۰.۹۵	۰.۹۶	۹۵
	برچسب ۰	۰.۹۶	۰.۹۵	۰.۹۸	۰.۹۷	۱۰٤
رگرسیون لجستیک	برچسب ۱	۰.۹۷	۰.۹۹	۰.۹۶	۰.۹۷	۹۵
	برچسب ۰	۰.۹۷	۰.۹۶	۰.۹۹	۰.۹۸	۱۰٤
تقویت گرادیان	برچسب ۱	۰.۹۵	۰.۹۷	۰.۹۳	۰.۹۵	۹۵
	برچسب ۰	۰.۹۵	۰.۹٤	۰.۹۷	۰.۹۵	۱۰٤
تقویت گرادیان شدید	برچسب ۱	۰.۹۶	۰.۹۸	۰.۹٤	۰.۹۶	۹۵
	برچسب ۰	۰.۹۵	۰.۹٤	۰.۹۸	۰.۹۶	۱۰٤

معیارهای ارزیابی انتخاب‌شده شامل صحت، دقت، یادآوری و F۱ به‌دلیل توان بالای آن‌ها در نمایش جزئیات عملکرد مدل در مسائل طبقه‌بندی دودویی انتخاب شدند. این معیارها امکان بررسی جامع‌تری از کارایی مدل در تشخیص نمونه‌های مثبت و منفی را فراهم می‌کنند. همچنین نمودار ROC برای مدل‌ها ترسیم شد، اما به‌دلیل شباهت زیاد منحنی‌ها و هم‌پوشانی بالا، امکان تحلیل دقیق تفاوت بین مدل‌ها از این طریق وجود نداشت. به همین دلیل تحلیل عملکرد بر مبنای معیارهایی که تمایز آشکارتری بین الگوریتم‌ها ایجاد می‌کردند انجام پذیرفت.



شکل ۳: ماتریس‌های سردرگمی الگوریتم‌های یادگیری ماشین: الف) درخت تصمیم، ب) جنگل تصادفی، ج) کا-نزدیک‌ترین همسایه، د) ماشین بردار پشتیبان، ه) رگرسیون لجستیک، و) تقویت گرادیان، ز) تقویت گرادیان شدید

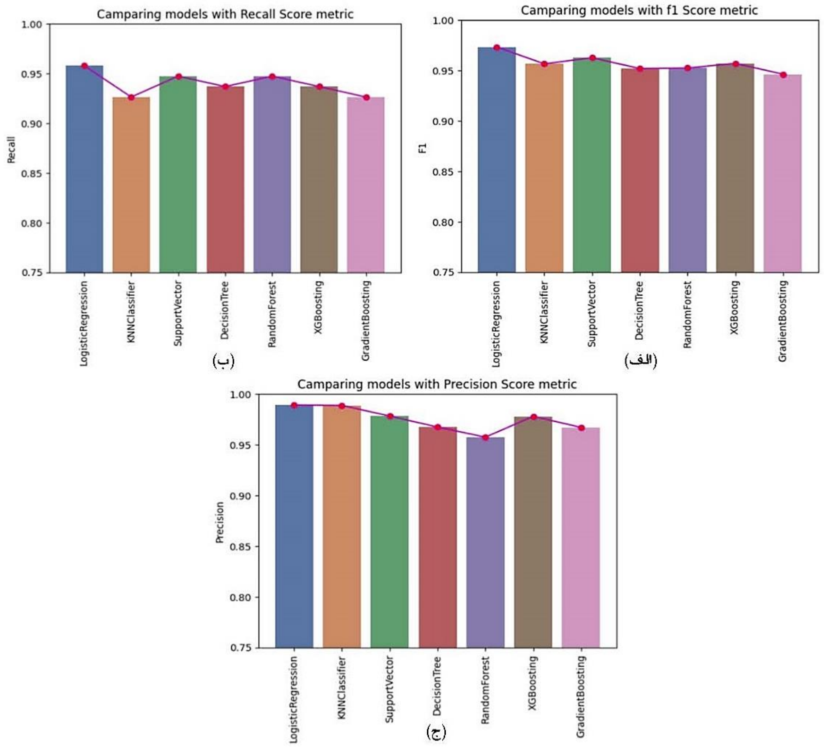
۴-۲- مقایسه مدل‌ها

در این بخش، ۷ الگوریتم یادگیری ماشین مورد بررسی قرار گرفتند و هر یک با استفاده از شاخص‌های صحت، دقت، یادآوری، F۱ و پشتیبانی ارزیابی شدند. از الگوریتم‌های جنگل تصادفی، درخت تصمیم، کا-نزدیک‌ترین همسایه، ماشین بردار پشتیبان، رگرسیون لجستیک و تقویت گرادیان برای ساخت مدل پایه پیش‌بینی استفاده شد. نتایج مقایسه این مدل‌ها نشان می‌دهد که برای این پژوهش، مناسب‌تر است که ابتدا از مدل رگرسیون لجستیک استفاده کنیم (رجوع کنید به جدول ٤ و اشکال ٤ و ٥).

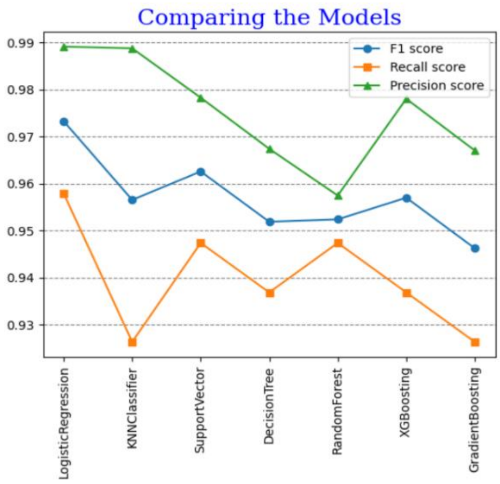
جدول ۴: مقایسه مدل‌ها بر حسب معیارهای ارزیابی

	دقت	یادآوری	F۱
رگرسیون لجستیک	۰.۹۸۹۱۳۰	۰.۹۵۷۸۹۵	۰.۹۷۳۲٦۲
کا-نزدیک‌ترین همسایه	۰.۹۸۸۷٦٤	۰.۹۲٦۳۱٦	۰.۹٥٦٥۲۲
ماشین بردار پشتیبان	۰.۹۷۸۲٦۱	۰.۹٤۷۳٦۸	۰.۹٦۲٥٦۷
درخت تصمیم	۰.۹٦۷۳۹۱	۰.۹۳٦۸٤۲	۰.۹٥۱۸۷۲
جنگل تصادفی	۰.۹٥۷٤٤۷	۰.۹٤۷۳٦۸	۰.۹٥۲۳۸۱
تقویت گرادیان شدید	۰.۹۷۸۰۲۲	۰.۹۳٦۸٤۲	۰.۹٥٦۹۸۹
تقویت گرادیان	۰.۹٦۷۰۳۳	۰.۹۲٦۳۱٦	۰.۹٤٦۲۳۷

با مقایسه جدول ٤ و اشکال ٤ و ٥ به این نتیجه خواهیم رسید که باتوجه‌به حجم کم مجموعه‌داده درنظر گرفته برای این پژوهش، تمامی مدل‌های پایه در نظر گرفته شده، عملکرد بسیار خوبی در پیش‌بینی دارند و استفاده از هر کدام از آن‌ها نتایج نسبتاً مشابهی به ما خواهد داد اما در این بین روش لجستیک رگرسیون با صحت ۰.۹٥، دقت ۰.۹۹، یادآوری ۰.۹٥، F۱ ۰.۹٦و پشتیبانی ۹٥ و همچنین روش ماشین بردار پشتیبان با صحت ۰.۹٦، دقت ۰.۹۸، یادآوری ۰.۹٥، F۱ ۰.۹٦و پشتیبانی ۹٥ عملکرد نسبتاً بهتری در پیش‌بینی داشته‌اند و ما این دو روش را به‌عنوان مدل پایه نسبت به مدل‌های دیگر به‌ویژه در حجم کلان داده‌ها برتری می‌دهیم.



شکل ۴: مقایسه مدل‌ها بر حسب معیارهای الف) F1 ب) یادآوری ج) صحت



شکل ۵: مقایسه مدل‌ها بر حسب سه معیار F1، یادآوری و صحت

۴-۳- پیش‌بینی نهایی

برای ترکیب نتایج مدل‌های پایه و بهبود دقت کلی، از رویکرد یادگیری جمعی مبتنی بر پشته‌سازی استفاده شد. در این چارچوب، رگرسیون لجستیک و ماشین بردار پشتیبان (SVM) به‌عنوان مدل‌های پایه (Level-۱) آموزش داده شدند. خروجی احتمالاتی حاصل از هر یک از این مدل‌ها استخراج و به‌عنوان ویژگی‌های ورودی به متامدل منتقل گردید. در سطح فرا–مدل (Level-۲)، یک رگرسیون لجستیک نهایی آموزش داده شد تا بر اساس این ویژگی‌ها پیش‌بینی قصد خرید را انجام دهد. آموزش مدل‌های پایه و متامدل بر روی ۸۰ درصد داده‌ها (مجموعه آموزشی) انجام گرفت و ارزیابی عملکرد نهایی روی ۲۰ درصد داده‌ها (مجموعه آزمون) صورت پذیرفت. این سازوکار موجب بهره‌گیری از نقاط قوت هر دو الگوریتم و افزایش دقت کلی پیش‌بینی شد.

۴-۴- ارزیابی نهایی مدل

پس از انجام تمامی مراحل پیاده‌سازی، نوبت به ارزیابی عملکرد مدل پیشنهادی می‌رسد که با استفاده از معیارهای رایج ارزیابی انجام شده است (به جدول ۵ مراجعه شود). نتایج نشان می‌دهد که الگوریتم عملکرد خوبی در پیش‌بینی قصد خرید کاربران دارد.

جدول ۵: نتیجه ارزیابی متامدل ساخته‌شده در این پژوهش

	صحت	دقت	یادآوری	F۱	پشتیبانی
برچسب ۱	۰.۹٦	۰.۹۸	۰.۹٥	۰.۹٦	۹٥
برچسب ۰	۰.۹٦	۰.۹٥	۰.۹۸	۰.۹۷	۱۰٤

این پژوهش به ارائه مدلی جدید برای پیش‌بینی قصد خرید پرداخت که در مقایسه با کارهای پیشین، نوآوری خاصی داشت. تمرکز مدل ما بر بازخورد کاربران به تبلیغات رسانه‌های اجتماعی بود، که در پژوهش‌های قبلی کمتر به آن توجه شده است. همچنین، استفاده از ویژگی‌های دموگرافیک برای پیش‌بینی قصد خرید، تنها به‌عنوان یک ایده در تحقیقات بیان شده بود و ما توانستیم آن را به مرحله اجرا درآوریم. در این راستا، علاوه بر داده‌های نشست سایت، از ویژگی‌هایی نظیر سن، جنسیت و میزان درآمد نیز استفاده کردیم..

۴-۵- مقایسه مدل

ما جهت سنجش عملکرد مدل پیشنهادی خود، آن را با سه مدل پیش‌بینی مورد استفاده در پژوهش‌های پیشین (جنگل تصادفی، درخت تصمیم، کا-نزدیک‌ترین همسایه) تحت شرایط یکسان آزمودیم. ما این ٤ مدل را با مجموعه‌داده توصیف شده در بخش و ویژگی‌های انتخاب-شده برای آن تحت ٥ معیار دقت، صحت، یادآوری، F۱ و پشتیبانی سنجیدیم. نتیجه در جدول ٦ قابل‌مشاهده است.

جدول ۶: نتیجه معیارهای ارزیابی برای هر یک از روش‌های پیش‌بینی قصد خرید موجود

	دقت	صحت	یادآوری	F۱	پشتیبانی
مدل پیشنهادی پژوهش ما	۰.۹٦	۰.۹۸	۰.۹٥	۰.۹٦	۹٥
جنگل تصادفی	۰.۹٥	۰.۹٦	۰.۹٦	۰.۹٥	۹٥
درخت تصمیم	۰.۹٥	۰.۹۷	۰.۹٤	۰.۹٥	۹٥
k-نزدیک‌ترین همسایه	۰.۹٦	۰.۹۹	۰.۹۳	۰.۹٦	۹٥

طبق جدول ٦ و همانطور که پیش‌تر گفته شد؛ تمامی مدل‌ها باتوجه‌به حجم کم دیتاهای مورد استفاده در این پژوهش، عملکرد تقریباً یکسانی دارند اما در حجم داده‌های بسیار بیشتر، مدل پیشنهادی ذکر شده از سایرین عملکرد بهتری دارد. همچنین در پژوهش کبیر و همکاران (۲۰۱۹)، دقت الگوریتم درخت تصمیم برابر با ۰.۸٦ و جنگل تصادفی برابر با ۰.۸۹ گزارش شده است [١٤]. در پژوهش حاضر، با پیاده‌سازی مدل ترکیبی مبتنی بر پشته‌سازی، دقت ۰.۹٦ به‌دست آمد که نشان‌دهنده بهبود عملکرد نسبت به مدل‌های پژوهش‌های پیشین است. نتایج نشان می‌دهند که روش ترکیبی پشته‌سازی، عملکرد برتری نسبت به سایر الگوریتم‌ها داشته‌است. این برتری را می‌توان به توانایی آن در یادگیری الگوهای پیچیده و کاهش بیش‌برازش نسبت داد. همچنین، روش پشته‌سازی با ترکیب قدرت چند الگوریتم، به تعادلی مناسب بین معیارهای دقت و F۱ دست یافته است. در مقابل، الگوریتم‌هایی مانند کا-نزدیکترین همسایه و درخت تصمیم به‌دلیل حساسیت به ساختار داده و ویژگی‌هایی با قدرت تفکیک پایین یا اطلاعات پیش‌بینی محدود، نتایج ضعیف‌تری ارائه کرده‌اند. این تفاوت‌ها به ماهیت ذاتی هر الگوریتم و شیوه پردازش ویژگی‌ها بازمی‌گردد.

۵- نتیجه‌گیری

رشد اینترنت و دیجیتالی‌شدن باعث افزایش محبوبیت خرید آنلاین شده و تحلیل قصد خرید مشتری به یکی از موضوعات کلیدی در بازاریابی دیجیتال تبدیل گردیده است [۱۳]. شرکت‌ها با تحلیل رفتار مشتری در این رسانه‌ها، اطلاعات ارزشمندی در مورد علایق آنان به دست می‌آورند [۲]. در این پژوهش، مدلی ترکیبی برای پیش‌بینی قصد خرید ارائه شد که بر پایه روش پشته‌سازی و ترکیب الگوریتم‌های رگرسیون لجستیک و ماشین بردار پشتیبان عمل می‌کند. در ابتدا از چند الگوریتم پایه یادگیری ماشین برای آموزش داده‌ها و پیش‌بینی اولیه استفاده شد. براساس ارزیابی‌ها، رگرسیون لجستیک و ماشین بردار پشتیبان عملکرد بهتری نسبت به سایر الگوریتم‌ها داشتند و خروجی آن‌ها به‌عنوان ورودی متامدل مورد استفاده قرار گرفت. یافته‌ها نشان دادند که ترکیب داده‌های رفتاری مرتبط با تبلیغات (مانند کلیک و مدت‌زمان حضور) با ویژگی‌های دموگرافیک کاربران از جمله سن، جنسیت و درآمد موجب افزایش دقت پیش‌بینی نسبت به مدل‌های منفرد می‌شود.

نوآوری اصلی پژوهش در تمرکز بر ادغام داده‌های رفتاری و جمعیت‌شناختی در چارچوبی ترکیبی بود که دقت بالاتری در پیش‌بینی قصد خرید فراهم ساخت. نتایج نشان داد که کاربران با مدت زمان حضور بالاتر در سایت و درآمد متوسط بیشترین نرخ کلیک را دارند. این موضوع می‌تواند برای طراحی تبلیغات هدفمند، بهینه‌سازی زمان نمایش تبلیغات و شخصی‌سازی پیام‌های بازاریابی به کار گرفته شود. همچنین پیاده‌سازی چنین مدلی در پلتفرم‌های تجارت الکترونیک می‌تواند نرخ تبدیل را افزایش دهد و بودجه تبلیغاتی را به‌صورت مؤثرتری تخصیص دهد. بدین ترتیب، پژوهش حاضر علاوه بر ارائه یک مدل علمی، کاربردهای عملی ارزشمندی برای مدیران بازاریابی دیجیتال نیز دارد.

در ارتباط با پرسش‌های اصلی تحقیق، نتایج نشان داد که ترکیب ویژگی‌های رفتاری کاربران (مانند کلیک و مدت‌زمان تعامل) با داده‌های دموگرافیک می‌تواند پیش‌بینی قصد خرید را به‌طور معناداری بهبود دهد. همچنین به‌کارگیری مدل ترکیبی مبتنی بر روش پشته‌سازی نسبت به مدل‌های منفرد عملکرد دقیق‌تر و پایاتری در پیش‌بینی قصد خرید ارائه داد. این نتایج تأییدی بر کارایی چارچوب پیشنهادی پژوهش هستند.

با وجود این دستاوردها، پژوهش حاضر دارای چند محدودیت است. نخست، اندازه مجموعه‌داده نسبتاً محدود (۱۰۰۰ نمونه) بوده که می‌تواند تعمیم‌پذیری نتایج را کاهش دهد. دوم، ارزیابی مدل تنها بر یک مجموعه‌داده خاص انجام شد و برای اطمینان از پایایی نتایج، از روش‌های اعتبارسنجی پیشرفته مانندCross-validation یا Bootstrap استفاده نشد. سوم، اگرچه از سه ویژگی دموگرافیک (سن، جنسیت و درآمد) بهره گرفته شد، اما متغیرهایی چون سطح تحصیلات، شغل یا سواد کامپیوتری لحاظ نشدند که می‌توانند بر قصد خرید اثرگذار باشند. چهارم، به دلیل محدودیت حجم مقاله، تحلیل‌های تکمیلی مانند اعتبارسنجی متقابل و تحلیل حساسیت ویژگی‌ها انجام نشد که می‌تواند در تحقیقات آینده مدنظر قرار گیرد.

برای گسترش پژوهش‌های آتی، مسیرهای متعددی وجود دارد. نخست، استفاده از مجموعه‌داده‌های بزرگ‌تر و متنوع‌تر می‌تواند به تعمیم‌پذیری و پایداری نتایج کمک کند. دوم، بهره‌گیری از داده‌های تکمیلی مانند تعاملات کاربران در شبکه‌های اجتماعی (نظرات، لایک‌ها و میزان اشتراک‌گذاری) می‌تواند ابعاد رفتاری بیشتری وارد مدل سازد. سوم، بررسی کاربران غیرخریدار و توسعه چارچوبی برای تحلیل رفتار آنان می‌تواند دیدگاه جامع‌تری نسبت به بازار فراهم آورد. چهارم، آزمون مدل در شرایط واقعی با داده‌های عملیاتی و کاربران واقعی وب‌سایت‌های خرید آنلاین می‌تواند اعتبار یافته‌ها را افزایش دهد. پنجم، استفاده از الگوریتم‌های پیچیده‌تر مانند شبکه‌های عصبی بازگشتی، حافظه طولانی‌کوتاه‌مدت (LSTM) و سایر مدل‌های یادگیری عمیق می‌تواند دقت و پایداری پیش‌بینی را بهبود بخشد. ششم، بررسی متغیرهایی چون نوع محصول مشاهده‌شده، موقعیت جغرافیایی کاربر و زمان تعامل با تبلیغات نیز می‌تواند مسیرهای ارزشمندی برای پژوهش‌های بعدی باشد.

۶- منابع

[۱] Asadollahi, M., S. R. S. Javadin and M. Nazari (۲۰۲٤). "Developing a Framework for Effective Digital Nudges in Online Shopping: A Systematic Literature Review." Management Research in Iran ۲۸(۳): ۱۷۹-۲۰۸. [In Persian]

[۲] Esmeli, R., M. Bader-El-Den and H. Abdullahi (۲۰۲۱). "Towards early purchase intention prediction in online session based retailing systems." Electronic Markets ۳۱(۳): ٦۹۷-۷۱٥. https://doi.org/۱۰,۱۰۰۷/s۱۲٥۲٥-۰۲۰-۰۰٤٤۸-x.

[۳] Shahrokh, Z. D. and M. Naeli (۲۰۲۰). "The Impact of Social Media Marketing Activities On Customer Equity of Luxury Brands A Study of Dorsa Brand." Management Research in Iran ۲٤(۱): ۳۳-۵۷. [In Persian]

[٤] Algawiaz, D., G. Dobbie and S. Alam (۲۰۱۹). Predicting a user's purchase intention using AdaBoost. ۲۰۱۹ IEEE ۱٤th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), IEEE. https://doi.org/۱۰,۱۱۰۹/ISKE٤۷۸۵۳,۲۰۱۹,۹۱۷۰۳۱٦.

[۵] Drechsler, W. and M. Natter (۲۰۱۱). "Do price charts provided by online shopbots influence price expectations and purchase timing decisions?" Journal of Interactive Marketing ۲۵(۲): ۹۵-۱۰۹. https://doi.org/۱۰,۱۰۱٦/j.intmar.۲۰۱۱,۰۲,۰۰۱.

[٦] Lamrhari, S., H. El Ghazi, M. Oubrich and A. El Faker (۲۰۲۲). "A social CRM analytic framework for improving customer retention, acquisition, and conversion." Technological Forecasting and Social Change ۱۷٤: ۱۲۱۲۷۵. https://doi.org/https://doi.org/۱۰,۱۰۱٦/j.techfore.۲۰۲۱,۱۲۱۲۷۵.

[۷] Nasir, V. A., A. C. Keserel, O. E. Surgit and M. Nalbant (۲۰۲۱). "Segmenting consumers based on social media advertising perceptions: How does purchase intention differ across segments?" Telematics and informatics ٦٤: ۱۰۱٦۸۷. https://doi.org/https://doi.org/۱۰,۱۰۱٦/j.tele.۲۰۲۱,۱۰۱٦۸۷.

[۸] Pavlou, P. A. (۲۰۰۳). "Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model." International journal of electronic commerce ۷(۳): ۱۰۱-۱۳٤. https://doi.org/۱۰,۱۰۸۰/۱۰۸٦٤٤۱۵,۲۰۰۳,۱۱۰٤٤۲۷۵.

[۹] Wu, P. C., G. Y.-Y. Yeh and C.-R. Hsiao (۲۰۱۱). "The effect of store image and service quality on brand image and purchase intention for private label brands." Australasian Marketing Journal ۱۹(۱): ۳۰-۳۹. https://doi.org/https://doi.org/۱۰,۱۰۱٦/j.ausmj.۲۰۱۰,۱۱,۰۰۱.

[۱۰] Fard, R. S., H. Hosseinzadeh and M. Khajeh (۲۰۲۳). "Deciding On The Main Components Affecting The Social Media Marketing Model By Meta-Synthesis." Journal of Modern Research in Decision Making ۸(۳): ۱۳۳-۱٦۱. [In Persian]

[۱۱] Wang, S.-C., Y.-T. Tsai and Y.-S. Ciou (۲۰۲۰). "A hybrid big data analytical approach for analyzing customer patterns through an integrated supply chain network." Journal of Industrial Information Integration ۲۰: ۱۰۰۱۷۷. https://doi.org/https://doi.org/۱۰,۱۰۱٦/j.jii.۲۰۲۰,۱۰۰۱۷۷.

[۱۲] Chu, Y., H.-K. Yang and W.-C. Peng (۲۰۱۹). Predicting online user purchase behavior based on browsing history. ۲۰۱۹ IEEE ۳۵th international conference on data engineering workshops (ICDEW), IEEE. https://doi.org/۱۰,۱۱۰۹/ICDEW.۲۰۱۹,۰۰-۱۳.

[۱۳] Bag, S., M. K. Tiwari and F. T. Chan (۲۰۱۹). "Predicting the consumer's purchase intention of durable goods: An attribute-level analysis." Journal of Business Research ۹٤: ٤۰۸-٤۱۹. https://doi.org/https://doi.org/۱۰,۱۰۱٦/j.jbusres.۲۰۱۷,۱۱,۰۳۱.

[۱٤] Kabir, M. R., F. B. Ashraf and R. Ajwad (۲۰۱۹). Analysis of different predicting model for online shoppers' purchase intention from empirical data. ۲۰۱۹ ۲۲nd International Conference on Computer and Information Technology (ICCIT), IEEE. https://doi.org/۱۰,۱۱۰۹/ICCIT٤٨٨٨٥,۲۰۱۹,۹۰۳۸٥۲۱.

[۱٥] Karthik, R. and S. Ganapathy (۲۰۲۱). "A fuzzy recommendation system for predicting the customers interests using sentiment analysis and ontology in e-commerce." Applied Soft Computing ۱۰۸: ۱۰۷۳۹٦. https://doi.org/۱۰,۱۰۱٦/J.ASOC.۲۰۲۱,۱۰۷۳۹٦.

[۱٦] Alalwan, A. A. (۲۰۱۸). "Investigating the impact of social media advertising features on customer purchase intention." International journal of information management ٤۲: ٦٥-۷۷. https://doi.org/https://doi.org/۱۰,۱۰۱٦/j.ijinfomgt.۲۰۱۸,۰٦,۰۰۱.

[۱۷] Noroozi, M., A. Khadivar and F. Abbasi (۲۰۲۳). "Modeling and predicting mobile phone purchase intention of Twitter users based on sentiment analysis." Journal of Modern Research in Decision Making ۸(۱): ۹۱-۱۱۲. [In Persian]

[۱۸] Sreejesh, S., J. Paul, C. Strong and J. Pius (۲۰۲۰). "Consumer response towards social media advertising: Effect of media interactivity, its conditions and the underlying mechanism." International Journal of Information Management ٥٤: ۱۰۲۱٥٥. https://doi.org/۱۰,۱۰۱٦/j.ijinfomgt.۲۰۲۰,۱۰۲۱٥٥.

[۱۹] Deng, Q., M. J. Hine, S. Ji and Y. Wang (۲۰۲۱). "Understanding consumer engagement with brand posts on social media: The effects of post linguistic styles." Electronic Commerce Research and Applications ٤۸: ۱۰۱۰٦۸. https://doi.org/https://doi.org/۱۰,۱۰۱٦/j.elerap.۲۰۲۱,۱۰۱۰٦۸.

[۲۰] Chang, Y., Y. Li, J. Yan and V. Kumar (۲۰۱۹). "Getting more likes: The impact of narrative person and brand image on customer–brand interactions." Journal of the Academy of Marketing Science ٤۷: ۱۰۲۷-۱۰٤٥. https://doi.org/۱۰,۱۰۰۷/s۱۱۷٤۷-۰۱۹-۰۰٦۳۲-۲.

[۲۱] Duong, H. T., A. A. Amaya Rivas and Y.-K. Liao (۲۰۱۹). "Examining the influence of customer-to-customer electronic word-of-mouth on purchase intention in social networking sites." https://doi.org/https://doi.org/۱۰,۱۰۱٦/j.apmrv.۲۰۱۸,۰٦,۰۰۳.

[۲۲] Patel, V., K. Das, R. Chatterjee and Y. Shukla (۲۰۲۰). "Does the interface quality of mobile shopping apps affect purchase intention? An empirical study." Australasian Marketing Journal ۲۸(٤): ۳۰۰-۳۰۹. https://doi.org/https://doi.org/۱۰,۱۰۱٦/j.ausmj.۲۰۲۰,۰۸,۰۰٤

[۲۳] McClure, C. and Y.-K. Seock (۲۰۲۰). "The role of involvement: Investigating the effect of brand's social media pages on consumer purchase intention." Journal of retailing and consumer services ٥۳: ۱۰۱۹۷٥. https://doi.org/https://doi.org/۱۰,۱۰۱٦/j.jretconser.۲۰۱۹,۱۰۱۹۷٥.

[۲٤] Bazi, S., R. Filieri and M. Gorton (۲۰۲۰). "Customers' motivation to engage with luxury brands on social media." Journal of Business Research ۱۱۲: ۲۲۳-۲۳٥. https://doi.org/https://doi.org/۱۰,۱۰۱٦/j.jbusres.۲۰۲۰,۰۲,۰۳۲.

[۲۵] Ullah, N., M. F. Gulzar and A. A. Arshad (۲۰۲۵). "Enhancing Consumer Purchase Intention in Pakistan Market: The Role of AI Personalization, Prediction Accuracy, and Real-Time Engagement." Asian American Research Letters Journal۲(۱): ۶۹-۷٤. https://doi.org/۱۰,۵۲۸۱/hc۸۱xc۵۳

[۲۶] Alsaadi, H., A. Wali and B. Fakieh (۲۰۲۵). "Predicting Impact of Digital Ads On Purchasing Intention Using Machine Learning Classification Models." Advances and Applications in Statistics ۹۲(۲): ۳۰۳-۳٤۲. https://doi.org/۱۰,۱۷٦٥٤/۰۹۷۲۳٦۱۷۲٥۰۱٤

[۲۷] Denis, J. C. and S. Baskaran (۲۰۲۵). "A predictive model for environmentally responsible purchase intention of organic products: a study from Malaysia." International Journal of Business Innovation and Research ۳٦(۲): ۲۰٤-۲۲۸. https://doi.org/۱۰,۱٥۰٤/IJBIR.۲۰۲٥,۱٤٤۳۹۳

[۲۸] Kotsiantis, S. B. (۲۰۱۳). "Decision trees: a recent overview." Artificial Intelligence Review ۳۹: ۲٦۱-۲۸۳. https://doi.org/۱۰,۱۰۰۷/s۱۰٤٦۲-۰۱۱-۹۲۷۲-٤.

[۲۹] Speiser, J. L., M. E. Miller, J. Tooze and E. Ip (۲۰۱۹). "A comparison of random forest variable selection methods for classification prediction modeling." Expert systems with applications ۱۳٤: ۹۳-۱۰۱. https://doi.org/https://doi.org/۱۰,۱۰۱٦/j.eswa.۲۰۱۹,۰٥,۰۲۸.

[۳۰] Guo, G., H. Wang, D. Bell, Y. Bi and K. Greer (۲۰۰۳). KNN model-based approach in classification. On The Move to Meaningful Internet Systems ۲۰۰۳: CoopIS, DOA, and ODBASE: OTM Confederated International Conferences, CoopIS, DOA, and ODBASE ۲۰۰۳, Catania, Sicily, Italy, November ۳-۷, ۲۰۰۳. Proceedings, Springer.

[۳۱] Vijayaragavan, P., R. Ponnusamy and M. Aramudhan (۲۰۲۰). "An optimal support vector machine based classification model for sentimental analysis of online product reviews." Future Generation Computer Systems ۱۱۱: ۲۳٤-۲٤۰. https://doi.org/۱۰,۱۱۰۹/BigData.۲۰۱۸,۸٦۲۲۲۳٥

[۳۲] Lv, Z., Y. Jin and J. Huang (۲۰۱۸). "How do sellers use live chat to influence consumer purchase decision in China?" Electronic Commerce Research and Applications ۲۸: ۱۰۲-۱۱۳. https://doi.org/۱۰,۱۰۱٦/j.elerap.۲۰۱۸,۰۱,۰۰۳

[۳۳] Shahbazi, Z., D. Hazra, S. Park and Y. C. Byun (۲۰۲۰). "Toward improving the prediction accuracy of product recommendation system using extreme gradient boosting and encoding approaches." Symmetry ۱۲(۹): ۱٥٦٦. https://doi.org/۱۰,۳۳۹۰/sym۱۲۰۹۱٥٦٦

[۳٤] XingFen, W., Y. Xiangbin and M. Yangchun (۲۰۱۸). Research on user consumption behavior prediction based on improved XGBoost algorithm. ۲۰۱۸ IEEE international conference on big data (Big Data), IEEE. https://doi.org/۱۰,۱۰۱٦/j.future.۲۰۲۰,۰٤,۰٤٦

[۲٤] Bazi, S., R. Filieri and M. Gorton (۲۰۲۰). "Customers' motivation to engage with luxury brands on social media." Journal of Business Research ۱۱۲: ۲۲۳-۲۳٥. https://doi.org/https://doi.org/۱۰,۱۰۱٦/j.jbusres.۲۰۲۰,۰۲,۰۳۲.

[۲٥] Ullah, N., M. F. Gulzar and A. A. Arshad (۲۰۲٥). "Enhancing Consumer Purchase Intention in Pakistan Market: The Role of AI Personalization, Prediction Accuracy, and Real-Time

Engagement." Asian American Research Letters Journal۲(۱): ٦۹-۷٤. https://doi.org/۱۰,٥۲۸۱/hc۸۱xc٥۳

[۲٦] Alsaadi, H., A. Wali and B. Fakieh (۲۰۲٥). "Predicting Impact of Digital Ads On Purchasing Intention Using Machine Learning Classification Models." Advances and Applications in Statistics ۹۲(۲): ۳۰۳-۳٤۲. https://doi.org/۱۰,۱۷٦٥٤/۰۹۷۲۳٦۱۷۲٥۰۱٤

[۲۷] Denis, J. C. and S. Baskaran (۲۰۲٥). "A predictive model for environmentally responsible purchase intention of organic products: a study from Malaysia." International Journal of Business Innovation and Research ۳٦(۲): ۲۰٤-۲۲۸. https://doi.org/۱۰,۱٥۰٤/IJBIR.۲۰۲٥,۱٤٤۳۹۳

[۲۸] Kotsiantis, S. B. (۲۰۱۳). "Decision trees: a recent overview." Artificial Intelligence Review ۳۹: ۲٦۱-۲۸۳. https://doi.org/۱۰,۱۰۰۷/s۱۰٤٦۲-۰۱۱-۹۲۷۲-٤.

[۲۹] Speiser, J. L., M. E. Miller, J. Tooze and E. Ip (۲۰۱۹). "A comparison of random forest variable selection methods for classification prediction modeling." Expert systems with applications ۱۳٤: ۹۳-۱۰۱. https://doi.org/https://doi.org/۱۰,۱۰۱٦/j.eswa.۲۰۱۹,۰٥,۰۲۸.

[۳۰] Guo, G., H. Wang, D. Bell, Y. Bi and K. Greer (۲۰۰۳). KNN model-based approach in classification. On The Move to Meaningful Internet Systems ۲۰۰۳: CoopIS, DOA, and ODBASE: OTM Confederated International Conferences, CoopIS, DOA, and ODBASE ۲۰۰۳, Catania, Sicily, Italy, November ۳-۷, ۲۰۰۳. Proceedings, Springer.

[۳۱] Vijayaragavan, P., R. Ponnusamy and M. Aramudhan (۲۰۲۰). "An optimal support vector machine based classification model for sentimental analysis of online product reviews." Future Generation Computer Systems ۱۱۱: ۲۳٤-۲٤۰. https://doi.org/۱۰,۱۱۰۹/BigData.۲۰۱۸,۸٦۲۲۲۳٥

[۳۲] Lv, Z., Y. Jin and J. Huang (۲۰۱۸). "How do sellers use live chat to influence consumer purchase decision in China?" Electronic Commerce Research and Applications ۲۸: ۱۰۲-۱۱۳. https://doi.org/۱۰,۱۰۱٦/j.elerap.۲۰۱۸,۰۱,۰۰۳

[۳۳] Shahbazi, Z., D. Hazra, S. Park and Y. C. Byun (۲۰۲۰). "Toward improving the prediction accuracy of product recommendation system using extreme gradient boosting and encoding approaches." Symmetry ۱۲(۹): ۱٥٦٦. https://doi.org/۱۰,۳۳۹۰/sym۱۲۰۹۱٥٦٦

[۳٤] XingFen, W., Y. Xiangbin and M. Yangchun (۲۰۱۸). Research on user consumption behavior prediction based on improved XGBoost algorithm. ۲۰۱۸ IEEE international conference on big data (Big Data), IEEE. https://doi.org/۱۰,۱۰۱٦/j.future.۲۰۲۰,۰٤,۰٤٦