مقایسه رویکرد مدل‌سازی پروبیت با لوجیت در اولویت‌بندی ویژگی‌های تاکسی اینترنتی: مقیاس‌بندی بهترین –بدترین نوع ۱

مقاله علمی – پژوهشی

محسن مکارمی شریفی*، دانش آموخته دکتری، گروه برنامه‌ریزی حمل‌ونقل، دانشکده فنی مهندسی، انشگاه بین‌المللی امام خمینی، قزوین، ایران

امیرعباس رصافی، استاد، گروه برنامه‌ریزی حمل‌ونقل، دانشکده فنی مهندسی، دانشگاه بین‌المللی امام خمینی، قزوین، ایران

*پست الکترونیکی نویسنده مسئول: m.makaremi.sh@gmail.com

دریافت: ۱۴۰۳/۰۳/۲۰– پذیرش: ۱۴۰۳/۰۹/۲۰

صفحه ۱۵۰-۱۳۷

چکیده

گسترش اینترنت و تحولات در دنیای کسب‌وکار، به ظهور کسب‌وکارهای جدید، از جمله تاکسی اینترنتی منجر شده است. این خدمت امکان سفارش تاکسی از طریق اپلیکیشن تلفن همراه را همراه با امکاناتی مانند پرداخت برخط و ارزیابی رانندگان فراهم می‌کند. رقابت در این زمینه، تاکسی‌های اینترنتی را به بهبود خدمات و جلب مشتریان بیشتر مثلا از طریق درک ترجیحات مشتریان و ارائه ویژگی‌هایی که بهترین تجربه را برای آنها فراهم می‌کند، سوق داده است. هدف این مطالعه بررسی این ویژگی‌ها (که در این‌جا ۱۰ ویژگی‌انتخاب شده) و اولویت‌بندی آنان است. در این مطالعه برخلاف شیوه سنتی پرداختن به مدل‌های انتخاب گسسته که تمرکز بر انتخاب بهترین (با اهمیت‌ترین) گزینه در میان انتخاب‌ها است، نقش بدترین (کم اهمیت‌ترین) گزینه نیز مورد توجه قرار می‌گیرد. از روش پروبیت برای مدل‌سازی استفاده شده و نتایج آن با روش لوجیت مورد مقایسه قرار گرفته‌اند. نتایج این پژوهش نشان می‌دهد که اگرچه ضرایب ویژگی‌ها در دو مدل‌سازی با هم یکسان نیستند و معیارهای ارزیابی برتری عملکرد مدل پروبیت را نسبت به مدل لوجیت نشان می‌دهد، ولی نتیجه کلی در رتبه‌بندی ویژگی‌ها با هم تفاوتی ندارند. در مواجهه با خدمات تاکسی اینترنتی، مسائل مرتبط با امنیت و اطمینان خاطر از اهمیت بیشتری برخوردار هستند. به عبارت دیگر کاربران در انتخاب خدمات تاکسی اینترنتی به جنبه‌هایی که احساس امنیت و رضایت ذهنی را برای آن‌ها به ارمغان می‌آورد، توجهبیشتری دارند. همچنین این مطالعه نشان می‌دهد که مسائل بهداشتی و فاصله‌گذاری اجتماعی در زمره کم‌اهمیت‌ترین ویژگی‌ها قرار دارند.

واژه‌های کلیدی: پروبیت، لوجیت، مقیاس‌بندی بهترین- بدترین نوع ۱، ویژگی‌های تاکسی اینترنتی

۱- مقدمه

آینده انسان به طور جدی تحت تأثیر اینترنت قرار گرفته و باعث تغییراتی چشمگیر در زندگی افراد شده است. از جمله کسب‌وکارهای موفقی که از این ارتباط گسترده بهره‌مند شده‌اند، خدمات تاکسی برخط است. اوبر، یکی از شرکت‌های پیشرو در این زمینه، در ایالات متحده آمریکا راه‌اندازی شد و تأثیرگذاری جهانی بسیار زیادی را به‌همراه داشت. در ایران نیز از حدود یک دهه قبل خدمات تاکسی برخط با معرفی "اسنپ" راه‌اندازی شد. در این مطالعه برای بررسی و مدل‌سازی انتخاب مسافران، از روش مقیاس‌بندی بهترین-بدترین نوع ۱ در مدل‌های انتخاب گسسته استفاده شده است. این رویکرد نه تنها به انتخاب بهترین گزینه توجه دارد، بلکه از انتخاب بدترین گزینه نیز در مدل‌سازی بهره می‌برد. استفاده از این رویکرد، درک پژوهشگر از ترجیحات مسافران و تأثیر ویژگی‌های مختلف بر انتخاب‌های آن‌ها را بهبود می‌دهد. این اطلاعات می‌تواند در بهبود تصمیم‌ها در زمینه سیستم‌های حمل‌ونقل همگانی و توسعه سیاست‌های مناسب برای تأثیر بر ترجیحات مسافران مؤثر باشد.

۲- تعریف مسأله و اهداف پژوهش

اهداف این مقاله بررسی ۱۰ ویژگی مهم در انتخاب خدمات تاکسی اینترنتی، اولویت‌بندی و تحلیل اهمیت و نقش هر یک از این ویژگی‌ها در انتخاب کاربران، مقایسه روش‌های مدل‌سازی پروبیت و لوجیت در تحلیل این ویژگی‌ها و نهایتا ارائه نتایجی است که برای بهبود خدمات تاکسی اینترنتی و افزایش رضایت کاربران مفید باشد. بر این اساس سؤالات زیر مطرح می‌شوند: چه ویژگی‌هایی در انتخاب خدمات تاکسی اینترنتی مهم به نظر می‌رسند؟ چگونه می‌توان این ویژگی‌ها را اولویت‌بندی و نقش هر یک را در فرآیند انتخاب تاکسی اینترنتی بررسی کرد؟ و آیا مدل پروبیت در مقایسه با مدل لوجیت نتایج بهتری در تحلیل ویژگی‌ها ارائه می‌دهد؟

۳- پیشینه تحقیق

هان رو لی با بررسی تجارب صنعتی، تفاوت‌های چشمگیر بین تاکسی‌ها و خودروهای ویژه را از نظر موقعیت و نحوه عملکرد در بازار مورد بررسی قرار داده و پیشنهادهایی را در زمینه تنظیم بازار، مدیریت عملیات و توسعه مشترک با خودروهای ویژه ارائه کرده است. وی بر این باور است که با توجه به اینکه صنعت تاکسی اصولاً مربوط به ارائه خدمات است و هدف نهایی آن بهبود کیفیت خدمات است، نوع شیوه‌های عملیاتی مورد استفاده تأثیر مستقیمی بر کیفیت ندارد(Li, 2016). داو و همکاران با ارائه یک سیستم مدیریت تاکسی مبتنی بر خدمات مکان‌محور و تکنیک‌های صف‌بندی منطقه‌ای در فضای اینترنت، بهبودهای چشمگیری را در این حوزه ارائه کرده‌اند. در این سیستم پیشنهادی مناطق خدماتی بر اساس ژئوگرید تعریف شده و به رانندگان این امکان را می‌دهند که هر دو روش شکار در خیابان‌ها و انتظار در ایستگاه‌ها را دنبال کنند (Dow et al., 2016).

کائو و لیو با هدف بهبود دقت در پیش‌بینی توزیع سفر تاکسی‌های خالی از طریق اپلیکیشن‌های تاکسی یک مفهوم نوین از تاکسی‌های خالی را بر اساس الگوی عملیات تاکسی مطرح کردند. آنها با استفاده از مدل لوجیت، توزیع سفر برای تاکسی‌های خالی و کل تاکسی‌ها را مدل کرده و نشان دادند که استفاده از اپلیکیشن‌های تاکسی به منظور کاهش تعداد سفرهای تاکسی‌های خالی دارای کاربردی عملی است (Yi & Xia, 2015). بومجاهد و همکاران به بررسی رفتار سفر در یک محیط بسیار متصل پرداخته‌اند و سعی کرده‌اند ابزارهای مبتنی بر فناوری اطلاعات و ارتباطات را برای تأثیرگذاری بر رفتارهای مسافران طراحی کنند. آنها به بررسی این پرسش پرداختند که چگونه پلتفرمی چون Yelp.com (یک شرکت چندملیتی درآمریکا که به مردم در یافتن کسب‌وکارهای محلی کمک می‌کند) می‌تواند به عنوان یک منبع اطلاعات برای فعالیت و برنامه‌ریزی سفر در مراحل پیش از سفر خدمت‌رسانی نماید(Mjahed et al., 2017).

با استفاده از داده‌های شبکه‌های اجتماعی بر اساس مکان برخط، چن و همکاران تأثیر شبکه‌های اجتماعی بر انتخاب مقصد مسافران در منطقه شهری شیکاگو را بررسی کردند. آنها از مدل‌های انتخاب سنتی برای پیش‌بینی تقاضای سفر استفاده کردند و فرض کردند که مجموعه انتخاب شناخته شده است و مسافران از ویژگی‌های گزینه‌های موجود آگاه هستند. نتایج این پژوهش نشان می‌دهد که روابط اجتماعی نقش مهمی در انتخاب مقصد مسافران دارند (Chen et al., 2018).

عوامل مؤثر بر پذیرش خدمات تاکسی اینترنتی در کالیفرنیا در میان نسل هزاره (متولدین سالهای ۱۹۸۱ تا ۱۹۹۷) و نسل قبل (متولدین سال‌های ۱۹۶۵ و ۱۹۸۰) توسط عالمی و همکاران بررسی شد و آنان نتیجه گرفتند که افراد تحصیل‌کرده و بزرگ‌تر نسل هزاره بیشتر احتمال دارد که خدمات تاکسی اینترنتی را استفاده کنند. همچنین کاربران همیشگی برنامه‌های گوشی‌های هوشمند مرتبط با حمل‌ونقل و کسانی که قبلا استفاده از تاکسی و خدمات خودروی اشتراکی را تجربه کرده‌اند، بیشتر احتمال دارد که از این خدمات استفاده کنند. این یافته‌ها می‌توانند به‌عنوان نقطه‌ای آغازین برای تحلیل و پیش‌بینی الگوهای رفتاری و تأثیرات اجتماعی و جمعیت‌شناسی در استفاده از خدمات حمل‌ونقل نوین در مناطق مختلف مورد استفاده قرار گیرند (Alemi et al., 2018). پژوهش‌ها در زمینه امنیت حمل‌ونقل برخط به‌ویژه در صنعت تاکسی اینترنتی نشان می‌دهد که امنیت یکی از مهمترین عوامل تأثیرگذار بر تصمیم‌های مسافران در استفاده یا استفاده مجدد از این سرویس‌ها است. در مطالعه‌ای که توسط پنگ جینگ و همکاران انجام شده است، به اهمیت اقدامات امنیتی در حمل‌ونقل برخط اشاره می‌کند و پیشنهاد می‌دهد که ارتقاء امنیت درک شده و کاهش خطر امنیتی می‌تواند تأثیر بسزایی در جلب اعتماد و استفاده افراد از خدمات تاکسی اینترنتی داشته باشد (Jing et al., 2021). در پژوهش انجام شده توسط اکبری و همکاران، عوامل مؤثر بر میزان استقبال کاربران از خدمات تاکسی اینترنتی در ایران مورد بررسی قرار گرفته‌اند. این پژوهش از ترکیب مدل پذیرش فناوری به همراه مدل موفقیت سیستم اطلاعات و

اعتماد بهره برده است. بر اساس این پژوهش، کیفیت خدمات و اطلاعات تأثیر چشمگیری بر سهولت استفاده و مطلوبیت درک‌شده دارد و مطلوبیت درک‌شده و سهولت استفاده درک‌شده هر دو با اعتماد به میزان مثبتی مرتبط بودند. مطلوبیت درک‌شده نیز با قصد رفتاری رابطه مثبتی داشت. با این حال رابطه مثبتی که بین سهولت استفاده درک‌شده و قصد رفتاری پیش‌بینی می‌شد، رد شد. علاوه بر این نتایج این پژوهش نقش اساسی اعتماد را به عنوان یک متغیر واسطه‌ای در مدل نشان داد (Akbari et al., 2020). رویکرد انتخاب گسسته پژوهش لوویر و وودورث نیز مقیاس‌بندی بهترین-بدترین بود. این رویکرد از اهمیت عوامل مختلف در فرآیند تصمیم‌گیری بهره می‌برد و افراد را قادر می‌سازد نه تنها بهترین گزینه را انتخاب کنند بلکه درک کنند که چه عواملی در این انتخاب تأثیرگذار بوده‌اند(Louviere & Woodworth, 1991). لوویر و فلین نیز تئوری و روش‌های مقیاس‌بندی بهترین-بدترین را گرد هم آوردند و کاربرد آنها را در مطالعات موردی مختلف در یک راهنمای مرجع مفید نشان دادند (Louviere et al., 2015).

هالیس و وستبری در زمینه پژوهش‌های روان‌زبانی، تفسیر و تحلیل کیفیت هنجارهای معنایی به‌دست‌آمده از مقیاس‌های مختلف را ارزیابی کردند و بررسی نتایج حاصله توسط آنان نشان می‌دهد که استفاده از مقیاس‌بندی بهترین- بدترین به‌طور کلی به تولید هنجارهایی با اعتبار پیش‌بینی بالاتر از سایر روش‌ها منجر می‌شود و در عین حال نیاز به داده کمتری برای این کار دارد. آنان همچنین نشان دادند که انتخاب قالب‌های مختلف پاسخ می‌تواند به داده‌های متنوع‌تری منجر شود که در تلاش‌های گذشته برای جمع‌آوری هنجارهای معنایی مورد توجه قرار نگرفته است(Hollis & Westbury, 2018).

لانکسار و همکاران یک راهنمای کاربر در مورد تجزیه و تحلیل داده‌های تولید شده از آزمایش‌های انتخاب گسسته و روش‌های بهترین-بدترین ارائه دادند. بر این اساس انتخاب روش مدل‌سازی به سؤالات و طراحی مطالعه و محدودیت‌ها از نظر کیفیت و کمیت داده‌ها بستگی دارد و تصمیم‌هایی که در رابطه با تجزیه‌وتحلیل داده‌های انتخاب گرفته می‌شوند، اغلب به جای ترتیبی بودن به طور متقابل به هم وابسته‌اند (Lancsar et al., 2017). با توجه به اهمیت حمل و نقل همگانی به ویژه در کشورهای در حال توسعه که بیشترین بخش سفرهای افراد جامعه تشکیل می‌دهند، والازا و همکاران در مطالعه‌ای تلاش کردند تا با استفاده از روش مقیاس‌بندی بهترین- بدترین، اولویت‌های مسافران را برای شهر کیپ‌تاون تشخیص دهند و براساس یافته‌های این پژوهش هرم کیفیت کاربر را ترسیم کنند. نتایج نشان داد که زمان سفر و هزینه آن اغلب به عنوان اولین اولویت‌ها در حمل و نقل همگانی منطقه‌ای مورد توجه قرار می‌گیرند. در این مطالعه، امنیت درون وسیله نقلیه و انتقال خدمات (تعداد دفعاتی که یک مسافر باید در یک سفر از یک وسیله نقلیه به وسیله نقلیه دیگر انتقال یابد) به عنوان ویژگی‌های کم اهمیت‌تر در کیفیت حمل و نقل همگانی شناخته شدند(Walaza et al., 2023).

پژوهشگران از جمله برتون و همکاران، مقیاس‌بندی بهترین-بدترین را به عنوان یک جایگزین برای رتبه‌بندی لیکرت در سنجش اولین برداشت چهره از شرکت‌کنندگان بررسی کردند. آنان به نتایجی دست یافتند که در سطح گروهی نمرات بهترین-بدترین تقریباً به طور کامل با نمرات لیکرت ارتباط دارد و نشان می‌دهد که این دو روش برداشت‌های یکسانی دارند. با این حال در سطح شرکت‌کنندگان فردی هم از نظر توانایی پیش‌بینی ترجیحات و هم از نظر آزمون اعتبارسنجی، رتبه‌بندی لیکرت پیشی می‌گیرد. این مزایا قدرت مقیاس‌بندی بهترین-بدترین را به‌ویژه برای استفاده در پژوهش‌های تفاوت‌های فردی ممتاز می‌سازد (Burton et al., 2019). در حوزه دیجیتال، بازاریابی محتوای دیجیتال که به عنوان یک مکمل برای رویکردهای ارتباطی بازاریابی سنتی در حال رشد و توسعه است، تنها زمانی تأثیر مثبت دارد که محتوای ارزشمندی برای مصرف‌کنندگان فراهم شود. کوب به منظور بررسی ترجیحات مصرف‌کنندگان از مقیاس‌بندی بهترین-بدترین استفاده کرد. وی با بررسی نظرات ۱۵۲۷ مصرف‌کننده از آلمان، سوئیس و اتریش نتیجه‌گیری کرد که مصرف‌کنندگان ارزش را به عنوان مهمترین مشخصه محتوا می‌پندارند و پس از آن ارزش در استفاده، ارزش سرگرمی، ارزش فرآیند و ارزش اجتماعی قرار دارد. این پژوهش همچنین نشان می‌دهد که مصرف‌کنندگان ویژگی‌های محتوای شرکت‌های دیجیتال را در وب‌سایت‌های شرکت متفاوت از آنچه در رسانه‌های اجتماعی انجام می‌دهند، اولویت‌بندی می‌کنند(Koob, 2023). وایت در مقاله‌ای بسته‌ی bwsTools یک مجموعه از ابزارهای رایگان و منبع باز برای زبان برنامه‌نویسی آماری R ارائه کرد که به محققان و متخصصان در ساخت و تجزیه و تحلیل طرح‌های مقیاس‌بندی بهترین-بدترین نوع ۱ کمک می‌نماید. این بسته برای کار یکپارچه با داده‌های مرتب طراحی شده است و نیازی به ماتریس‌های طراحی ندارد. همچنین از روش‌های مختلف امتیازدهی منتشر شده در سطح فردی و تجمیعی استفاده می‌کند که در نرم‌افزار آزاد استفاده نشده است(White, 2021).

آیزاکی در پژوهش‌های خود به نیاز به استفاده از بسته‌های نرم‌افزاری در محیط R اشاره کرده و با ذکر مثال‌هایی در حوزه مقیاس‌بندی بهترین-بدترین در مقالات خود توضیح داده است. همچنین در مقاله‌های اخیر خود توابع جدید و بازنگری شده مرتبط با این بسته‌ها را معرفی و ارائه کرده است (Aizaki & Fogarty, 2023).

٤-روش‌شناسی

٤-١-آزمایش‌های انتخاب گسسته (آ.ا.گ.)

آ.ا.گ. یک روش رجحان بیان شده است که شامل تولید و تجزیه و تحلیل داده‌های انتخابی است. آنها معمولاً در نظرسنجی‌ها اجرا می‌شوند. پاسخ‌دهندگان در این نظر سنجی در قالب چند سؤال با چند مجموعه انتخاب مواجه می‌شوند که هر یک شامل تعدادی گزینه است و از پاسخ‌دهندگان خواسته می‌شود تا از بین گزینه‌ها انتخاب کنند.

٤-٢- مدل انتخاب گسسته

مدل‌سازی انتخاب گسسته به عنوان یک روش تحلیلی اساسی برای درک بهتر رفتار تصمیم‌گیری مورد استفاده قرار می‌گیرد. این مدل‌سازی به پژوهشگران کمک می‌کند تا الگوهای مختلف تصمیم‌گیری را در میان داده‌های پیچیده آ.ا.گ. شناسایی کنند و اثرات مختلف ویژگی‌ها را بر ترجیحات تصمیم‌گیرندگان ارزیابی کنند. پژوهشگران با استفاده از مدل‌های انتخاب گسسته می‌توانند تأثیر هر ویژگی را بر انتخاب نهایی مورد بررسی قرار دهند و عوامل مؤثر در فرآیند تصمیم‌گیری را شناسایی کنند. این اطلاعات اساسی برای ارتقاء فهم ما از ترجیحات مصرف‌کنندگان و ارائه‌دهندگان خدمات است و در ارتقاء کیفیت خدمات و بهبود استراتژی‌های بازاریابی و تصمیم‌گیری کمک می‌کند. در آ.ا.گ. تصمیم گیرندگان به سادگی گزینه مورد علاقه خود را از میان مجموعه ای از گزینه‌ها در مجموعه‌های انتخابی انتخاب می‌کنند. مطلوبیتی را فرض کنید که پاسخ‌دهنده i از انتخاب گزینه j در سناریوی انتخابی s به دست می‌آورد. این مطلوبیت با رابطه (۱) نشان داده می‌شود.

$$U_{isj} = V_{isj} + \varepsilon_{isj} \qquad i= 1,...,N;\ s= 1,...,S;\ j=1,...,J \qquad (۱)$$

که در آن U_{isj} مطلوبیت واقعی گزینه n برای تصمیم‌گیرنده j و V_{isj} نشان‌دهنده مؤلفه سیستماتیک یا قابل پیش‌بینی مطلوبیت است. ε_{isj} خطا یا بخشی از مطلوبیت است که برای تحلیل‌گر ناشناخته می‌باشد. در رابطه (۱)، N تصمیم‌گیرنده در بین S سناریو از بین J گزینه انتخاب می‌کنند. $yis= j$ به این معناست که تصمیم‌گیرنده i گزینه j را از سناریوی s انتخاب می‌کند (در صورتی که بالاترین مطلوبیت را در مقایسه با مطلوبیت مرتبط با سایر گزینه‌های موجود در مجموعه انتخاب ارائه دهد). بنابراین احتمال انتخاب گزینه j به‌صورت زیر مدل می‌شود:

$$P_{isj}=Prob(y_{is}=j) = Prob(U_{isj} - U_{isl} > 0) \qquad \forall l \neq j \qquad (۲)$$

در این مطالعه رگرسیون لوجیت یا پروبیت که مدل لوجیت یا پروبیت نیز نامیده می‌شود، برای مدل‌سازی کمیت‌های وابسته دوگانه یا باینری استفاده می‌شود.

٤-٢-١- پروبیت

مدل‌های پروبیت ابزارهای آماری متداولی هستند که به طور معمول برای مدل‌سازی انتخاب گسسته استفاده می‌شوند. این مدل‌ها به ویژه زمانی مفید هستند که متغیر وابسته دودویی است (مانند انتخاب بین ماشین و حمل و نقل عمومی) یا ترتیبی است (مانند ترتیب اولویت‌ها برای حالت‌های مختلف). ایده اصلی پشت مدل‌های پروبیت در تخمین احتمال انتخاب یک شیوه خاص توسط یک فرد بر اساس مجموعه‌ای از متغیرهای توضیحی (مانند زمان سفر، هزینه، عوامل جمعیت‌شناختی) قرار دارد. برخلاف رگرسیون خطی، مدل‌های پروبیت فرض می‌کنند که بخش خطا (غیرقابل مشاهده) مرتبط با هر شیوه از توزیع نرمال استاندارد پیروی می‌کند. بر خلاف لوجیت برای این نوع مدل (به جز در حالات خاص) رابطه به شکل بسته وجود ندارد (Choe et al., 2020).

۲-۲-٤- لوجیت

مدل لوجیت با استفاده از تابع لوجیت که یک تابع لگاریتمی است، بر اساس فرض استقلال گزینه‌های نامرتبط (Independence of Irrelevant Alternatives) توسعه یافته است که به معنای این است که تغییر در یک گروه از داده‌ها نباید به تغییر در احتمال گروه دیگری منجر شود. فرض دیگر آن است که خطاها از توزیع گامبل پیروی می‌کنند (Choe et al., 2020). احتمال انتخاب j با استفاده از مدل لوجیت به شکل رابطه ۳ است.

$$P_{isj}=\frac{\exp(\lambda V_{isj})}{\sum_{l=1}^{J}\exp(\lambda V_{isl})} \quad (۳)$$

P_{isj}: احتمال انتخاب ویژگی j از سناریوی s برای فرد i،

V_{isj}: تابع مطلوبیت ویژگی j از سناریوی s برای فرد i، و

V_{jsl} : تابع مطلوبیت سایر ویژگی‌ها از سناریوی s برای فرد i.

λ پارامتری را نشان می‌دهد که مقیاس مطلوبیت را تعیین می‌کند، که متناسب با معکوس توزیع عبارات خطا است.

٤-۳- معیارهای ارزیابی مدل‌ها

ارزیابی مدل شامل گستره‌ای از روش‌هاست که هدف‌شان سنجش عملکرد و کیفیت مدل‌ها است. از بین معیارهای فراوانی که وجود دارند، چند نمونه از معروفترین آنها انتخاب شده‌اند تا عملکرد دو مدل پروبیت و لوجیت را بررسی کنیم.

٤-۳-۱- میانگین مربعات خطا یا MSE (Mean Squared Error)

میانگین مربعات خطا یا MSE یک معیار پرکاربرد در ارزیابی عملکرد مدل‌های پیش‌بینی و تخمینی است.

برای محاسبه MSE، ابتدا باید خطای پیش‌بینی مدل برای هر نمونه را محاسبه کنیم. این خطا معمولاً به عنوان تفاضل بین مقدار واقعی و مقدار پیش‌بینی شده برای هر نمونه تعریف می‌شود. سپس میانگین مربعات این خطاها برای تمام نمونه‌ها محاسبه می‌شود. فرمول محاسبه MSE به صورت زیر است.

$$\mathrm{MSE}=\frac{1}{n}\sum_{i=1}^{n}(yi-\hat{y}i)^2 \quad (٤)$$

در اینجا:

n تعداد کل نمونه‌ها است.

yi مقدار واقعی برای نمونه i ام است.

$\hat{y}i$ مقدار پیش‌بینی شده توسط مدل برای نمونه i ام است.

$(yi-\hat{y}i)^2$ مربع تفاضل بین مقدار واقعی و پیش‌بینی شده برای نمونه i ام است.

MSE با مربع گرفتن اختلافات بین مقادیر واقعی و پیش‌بینی شده، همه اختلافات را مثبت می‌کند و سپس میانگین می‌گیرد. این امر به معنای این است که اختلافات بزرگتر بیشتر تأثیر می‌گذارند و اختلافات کوچک‌تر تأثیر کمتری دارند.

MSE فرمول ساده و قابل فهم و قابلیت تفسیر راحت‌تری دارد. به طور کلی، مقدار MSE کمتر نشان‌دهنده عملکرد بهتر مدل در پیش‌بینی داده‌ها است.

٤-۳-۲- معیار AIC (Akaike Information Criterion)

AIC یک معیار است که برای مقایسه مدل‌های آماری مختلف استفاده می‌شود. AIC با استفاده از تعداد پارامترها و مقدار لگاریتم درست‌نمایی مدل، یک ارزیابی جامع از مدل ارائه می‌دهد. فرمول معیار AIC به صورت۵ است.

$$AIC=2k-2\ln(L) \quad (۵)$$

k تعداد پارامترهای مدل است.

L مقدار لگاریتم درست‌نمایی مدل است.

هدف از استفاده از معیار AIC، انتخاب مدلی است که هماهنگی خوبی با داده‌ها دارد، اما پارامترهای آن نیز ساده است. به عبارت دیگر، مدلی که AIC کمتری داشته باشد، بهتر است.

۴-۳-۳- ماتریس درهم ریختگی و معیار دقت

ماتریس خطا نیز شناخته می‌شود، یک جدول خلاصه‌شده برای ارزیابی عملکرد یک مدل طبقه‌بندی است. این معیار ارزیابی کیفیت عملکرد مدل دسته‌بندی را با استفاده از ماتریسی به نام ماتریس درهم ریختگی انجام می‌دهد. این ماتریس یک ماتریس مستطیلی است که در ستون‌ها و ردیف‌های خود دسته‌بندی‌های مختلف ممکن را شامل می‌شود و به صورت جدول (۱) است.

جدول ۱. مقایسه معیارهای ارزیابی در دو مدل لوجیت و پروبیت

		برآورد	
		غلط	درست
مشاهده	غلط	درست منفی (TN)	علط مثبت (FP)
	درست	غلط منفی (FN)	درست مثبت (TP)

TP (True Positive): تعداد نمونه‌های درستی که توسط مدل نیز به عنوان درست شناسایی شده‌اند.

FP (False Positive): خطای نوع ۱ نیز نامیده می‌شود و عبارت‌است از تعداد نمونه‌های غلط که به اشتباه توسط مدل به عنوان درست شناسایی شده‌اند.

TN (True Negative): تعداد نمونه‌های غلط که توسط مدل نیز به عنوان غلط شناسایی شده‌اند.

FN (False Negative): یک خطای نوع ۲ نیز نامیده می‌شود و عبارت‌است از تعداد نمونه‌های درستی که به اشتباه توسط مدل به عنوان غلط شناسایی شده‌اند.

معیار دقت یا Accuracy یکی از معیارهای مهم برای ارزیابی عملکرد مدل‌های دسته‌بندی است. این معیار به ما نشان می‌دهد که توانایی مدل در تشخیص صحیح داده‌ها چقدر است.

در مسائل دسته‌بندی، داده‌ها به چند دسته یا کلاس تقسیم می‌شوند و مدل سعی می‌کند هر داده را به درستی به یک کلاس اختصاص دهد. Accuracy ساده‌ترین روش برای اندازه‌گیری عملکرد یک مدل دسته‌بندی است و به این صورت محاسبه می‌شود:

$$\text{Accuracy} = \frac{\text{تعداد داده‌هایی که به درستی دسته‌بندی شده‌اند}}{\text{تعداد کل داده‌ها}} \quad (۶)$$

به طور معادل، می‌توان این را به صورت زیر نیز نوشت:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (۷)$$

در اینجا تعداد داده‌هایی که به درستی دسته‌بندی شده‌اند به معنای تعداد داده‌هایی است که مدل به درستی کلاس آن‌ها را تشخیص داده است و تعداد کل داده‌ها تعداد کل داده‌های موجود در مجموعه آزمون یا ارزیابی است. به عبارت دیگر دقت نشان می‌دهد که چه تعداد از تمام داده‌ها توسط مدل به درستی تشخیص داده شده‌اند. هرچه مقدار دقت بیشتر و به ۱ نزدیکتر باشد، عملکرد آن مدل بهتر است. دقت یک معیار بسیار ساده و قابل فهم و با قابلیت تفسیر راحت است که به طور مستقیم تعداد داده‌هایی را که به درستی دسته‌بندی شده‌اند نسبت به تعداد کل داده‌ها می‌سنجد.

٤-٤- مقیاس‌بندی بهترین-بدترین نوع ۱

مقیاس‌بندی بهترین-بدترین (م.ب.ب) یک روش مفید برای بررسی اولویت‌بندی و ارزش‌گذاری ویژگی‌ها و گزینه‌ها توسط افراد است. با استفاده از این مقیاس زیرمجموعه‌های مختلف از ویژگی‌ها ایجاد می‌شود و سپس با درخواست از افراد برای انتخاب بهترین و بدترین گزینه‌ها در هر زیرمجموعه، امکان مقایسه و ارزیابی اهمیت نسبی آنها فراهم می‌شود. این سوال می‌تواند تکرار شود تا زمانی که همه زیرمجموعه‌های یک مجموعه انتخاب ارزیابی شوند.

م.ب.ب. شامل سه نوع است که از نظر پیچیده بودن موارد یا گزینه‌های مورد بررسی متفاوت هستند. نوع۱ نیاز به فهرستی از اشیا (یا ویژگی‌های بدون سطوح) دارد که فرد می‌خواهد آن‌ها را اندازه‌گیری کند، و تصمیم‌گیرندگان باید بهترین و بدترین اشیاء را در مجموعه انتخاب برگزینند.

از آنجایی که چارچوب نظری روش م.ب.ب. شبیه به آ.ا.گ. است، نظریه مطلوبیت تصادفی از طریق مدل maxdiff (بیشینه تفاضل) در روش م.ب.ب. ارائه شده است. در مدل maxdiff، بهترین و بدترین انتخاب بالقوه به عنوان یک جفت تعریف می‌شوند (Guo & Shen, 2019).

٤-٥- تعیین متغیرها (ویژگی‌های تاکسی اینترنتی)

با مطالعه در پژوهش‌های پیشین و مصاحبه با متخصصین امر ۱۰ ویژگی بارز در این مقوله شناسایی شدند. این ویژگی‌ها در جدول (۲) آمده است.

جدول ۲. ویژگی‌های بارز تاکسی اینترنتی

هزینه	رعایت مسائل بهداشتی و فاصله‌گذاری اجتماعی
راحتی	سریع بودن
ایمنی در مقابل خطرات	سهولت دسترسی
امنیت و اطمینان خاطر	انعطاف‌پذیری (از نظر زمان و انتخاب مقصد واسطه و...)
تکریم مشتری	وابستگی به فناوری (گوشی و اینترنت)

٤-٦- شناسایی و تعیین محدوده مورد مطالعه

محدوده مورد مطالعه در این مقاله، استان قزوین و مرکز آن یعنی شهر قزوین است. استان قزوین در مرکز ایران واقع شده و مساحتی حدود ۱۵۶۲۳ کیلومتر مربع دارد. این استان از نظر جغرافیایی در محدوده ٤٨ درجه و ٤٤ دقیقه تا ٥٠ درجه و ٥١ دقیقه طول شرقی از نصف النهار گرینویچ و از عرض ۳۵ درجه و ٢٤ دقیقه تا ٣٦ درجه و ٤٨ دقیقه شمالی نسبت به خط استوا واقع شده است.

استان قزوین از سمت شمال به استان‌های گیلان و مازندران، از سمت غرب به استان‌های همدان و زنجان، از سمت جنوب به استان مرکزی و از سمت شرق به استان البرز محدود می‌شود. این استان شامل ۲۰ شهر در ۵ شهرستان است. به تازگی، تخمین جمعیت این استان تا پایان سال ۱۴۰۰ توسط مرکز ملی آمار ایران، یک میلیون و ۳۲٦ هزار و ٤٠٠ نفر اعلام شده است. همچنین جمعیت شهرستان قزوین برآورد شده به حدود ٦٢١ هزار و ۸۰۰ نفر می‌رسد.

٤-٧- انتخاب روش نمونه‌گیری و اندازه نمونه

علاوه بر اطلاعاتی که از روش کتابخانه‌ای به دست آمد، داده‌های مورد نیاز برای ساخت مدل نیز به روش میدانی جمع آوری شد. با توجه به چارچوب این رساله، به‌منظور جمع‌آوری اطلاعات مورد نیاز پرسشنامه‌هایی طراحی و بصورت الکترونیکی تهیه و لینک آن در شبکه‌های اجتماعی که افراد عضو آنها از طیف‌های مختلف سنی بودند، توزیع شد. پرسشنامه شامل سؤالاتی در مورد ویژگی‌های فردی (سن، جنس، تحصیلات، شغل و ...) و سؤالاتی نیز به منظور کشف

اولویت‌های بهترین-بدترین در میان گزینه‌ها و ویژگی‌ها مطرح شد. این روش جمع‌آوری داده‌ها از طریق شبکه‌های اجتماعی نه تنها امکان دسترسی به یک نمونه گسترده از افراد را فراهم کرد، بلکه به دلیل الکترونیکی بودن پرسشنامه‌ها امکان پردازش و تحلیل سریع‌تر داده‌ها نیز افزایش یافت. برای برآورد سرانگشتی حجم نمونه مورد نیاز برای مطالعات آ.ا.گ. با استفاده از یک یا چند ویژگی طراحی فرمول‌هایی ارائه شده که اولین آن را اورمه (۱۹۹۸) پیشنهاد کرد. بر اساس فرمول اورمه که در آن $Lmax$ بزرگترین تعداد سطوح برای هر یک از ویژگی‌ها است، J تعداد گزینه‌ها در هر انتخاب (بدون احتساب "هیچکدام") است، و S تعداد مجموعه‌های انتخاب یا بعبارتی تعداد سؤالاتی است که هر پاسخ دهنده به آن پاسخ می‌دهد(Assele et al., 2023).

$$N \geq 500\frac{Lmax}{JS} \quad (۸)$$

در این مطالعه حداکثر سطوح ویژگی‌ها ($Lmax$) برابر با ۳، تعداد گزینه‌ها در هر سوال (J) برابر با ٤ و تعداد سؤالات برای هر نفر(S) برابر با ۹ و بنابراین حداقل حجم نمونه ٤۲ نفر است. با توجه به اینکه مقیاس‌بندی بهترین-بدترین یک انتخاب متفاوت با نتایج متفاوت نسبت به نتایج آ.ا.گ.های سنتی را نشان می‌دهد، اندازه نمونه برای تخمین مطلوبیت‌ها ناشناخته است و هیچ روش اختصاصی برای تعیین اندازه نمونه در مورد م.ب.ب. مطرح نشده است. با این حال طبق نظر فلین و همکاران می‌توان از معادلات فواصل اطمینان برای برآورد اندازه نمونه‌های مورد نیاز بهره برد (Flynn et al., 2007). در این موارد مطابق نظر لوویر و همکاران معیارهای م.ب.ب. از قوانین مربوط به اندازه‌های نمونه مرتبط با توزیع‌های چندجمله‌ای پیروی می‌کنند (Louviere et al., 2013).

تامپسون (۱۹۸۷) فرمولی را برای محاسبه نیازهای اندازه نمونه چندجمله‌ای استخراج کرد. وی نشان داد که مانند حالت دوجمله‌ای، حجم نمونه تابعی از سطح خطای قابل قبول(α) و درجه اطمینان مطلوب مورد نیاز تحلیلگران در به دست آوردن نسبت‌های واقعی جمعیت (d) است. او همچنین نشان داد که اندازه نمونه مورد نیاز برای داده‌های نسبت چندجمله‌ای مستقل از تعداد دسته‌های چندجمله‌ای است (Thompson, 1987). اگر بدون درنظرگرفتن ملاحظات م.ب.ب. مایل به برآورد حجم نمونه باشیم، با توجه به آخرین برآورد نفوس و مسکن که جمعیت شهر قزوین را حدود ٦۰۰ هزار نفر برآورد کرده است، با استفاده از جدول مورگان و فرمول کوکران نیاز به ۳۸٤ مشاهده است(Ahmad & Halim, 2017). در این مطالعه ۱۰۰ پرسش‌نامه جمع‌آوری شد که در هر پرسش‌نامه ۱۲ سناریوی مختلف مورد پرسش قرار گرفت و در مجموع ۱۲۰۰ مشاهده حاصل شد. بنابراین با این تعداد مشاهده اطمینان خاطر از حجم نمونه مورد نیاز فراهم شد.

٤-۸- طراحی آزمایش طرح‌های تأثیرات اصلی متعامد

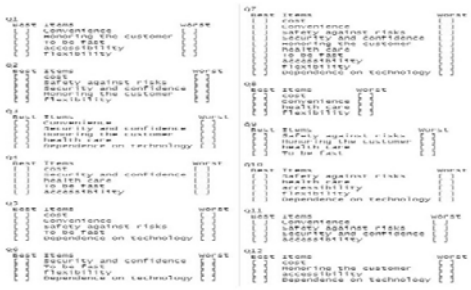
در این مطالعه از طراحی آزمایش طرح‌های تأثیرات اصلی متعامد (OMED) استفاده شده است. به این منظور در طراحی پرسش‌نامه با بهره‌گیری از تابع oa.design از پکیج DoE.base نرم‌افزار R طراحی آزمایش مورد نظر ساخته شد.

٤-۹- طراحی پرسش‌نامه

با استفاده از طراحی آزمایش طرح‌های تأثیرات اصلی متعامد (OMED) مجموعه انتخاب‌ها برای سوال از پاسخ‌دهندگان ساخته شد که در هر سؤال آن مجموعه‌ای از چند ویژگی ارائه و به صورت همزمان در معرض انتخاب با اهمیت ترین و کم اهمیت ترین گزینه برای انتخاب قرار گرفت. در این طراحی آزمایش از هر پاسخ‌دهنده ۱۲ سؤال پرسش شد که در هر سؤال ویژگی‌ها مطابق شکل (۱) به ترتیب زیر مطرح شد.

٤-۱۰- آمارگیری و جمع‌آوری داده‌ها

پس از طراحی سؤالات پرسش‌نامه، فرم نظرسنجی به صورت فرم برخط ساخته شد و لینک آن در اختیار ۱۰۰ نفر پاسخ‌دهنده قرار گرفت. پس از تکمیل نظرسنجی، داده‌های پاسخ که عبارت از انتخاب پاسخ‌دهندگان به عنوان بهترین و بدترین گزینه از هر مجموعه انتخاب است به داده‌های قابل استفاده جهت مدل‌سازی در نرم‌افزار R تبدیل می‌شوند.



شکل ۱. طراحی سؤالات پرسش‌نامه با استفاده از طراحی آزمایش طرح‌های تأثیرات اصلی متعامد

٤-١١- رویکرد مدل‌سازی در تجزیه و تحلیل

همان‌طور که پیشتر گفته شد از مدل‌های انتخاب گسسته برای تحلیل پاسخ‌ها استفاده می‌شود و مبتنی بر درک پاسخ‌های پاسخ‌دهندگان در شرایط زیر است: فرض کنید m مورد در یک مجموعه انتخاب (یک سوال) وجود دارد. تعداد جفت‌های ممکن که در آنها مورد i به‌عنوان بهترین و مورد j به‌عنوان بدترین ($i \neq j$) از m مورد انتخاب می‌شود $m\times (m-1)$

(۹)

شایان ذکر است که مخرج شامل مجموع تمام تفاوت‌های مطلوبیت بین موارد k و l است ($l \neq k$). اگر تعداد سؤالات در هر سؤال چهار سؤال باشد، تعداد بهترین-بدترین جفت‌های ممکن از دو سؤال از چهار سؤال، دوازده است. بنابراین ۱۲

(۱۰)

یافته‌های پژوهش

همانطور که قبلا بیان شد این پژوهش با پرسش از ۱۰۰ نفر انجام گرفت که ۵۱ نفر از آنان مرد و ٤۹ نفر زن بودند. ٤۰ نفر

تحلیل داده‌ها

با تحلیل داده‌ها در نرم‌افزار R، و با مدل‌سازی پروبیت و لوجیت با این فرض که ضریب ITEM8 (سهولت دسترسی) است. برای هر مورد، یک مطلوبیت (V) در نظر گرفته می‌شود که می‌تواند به دو بخش تقسیم شود: مؤلفه‌های سیستماتیک و تصادفی. فرض کنید که پاسخ‌دهندگان موارد i و j را به ترتیب به عنوان بهترین و بدترین انتخاب می‌کنند. تفاوت در مطلوبیت بین i و j نشان دهنده بزرگترین تفاوت مطلوبیت است. بر اساس این مفروضات، احتمال انتخاب مورد i به عنوان بهترین و مورد j به عنوان بدترین، در مدل لوجیت به صورت زیر بیان می‌شود.

$$P_r(i,j)=\frac{\exp(V_i-V_j)}{\sum_{k=1}^{m}\sum_{l=1,l\neq k}^{m}\exp(V_k-V_l)}$$

جفت مطلوبیت باید محاسبه شود. برای هر مطلوبیت برآورد شده، ضریب SPi "سهم ترجیحی" برای مورد i بصورت زیر محاسبه می‌شود:

$$SP_i=\frac{\exp(V_i)}{\sum_{j=1}^{m}\exp(V_j)}$$

از پرسش شوندگان کمتر از ۳۵ سال، ۳٦ نفر بین ۳۵ تا ۵۰ سال و ١٤ نفر نیز بیش از ۵۰ سال سن داشتند.

صفر درنظر گرفته شده است، نتایج شکل‌های ۲ و ۳ به دست می‌آیند.

ITEM1	ITEM2	ITEM3	ITEM4	ITEM5	ITEM6	ITEM7	ITEM8	ITEM9	ITEM10
-0.3487593	-0.3410433	-0.2221189	0.1393680	-0.4088033	-0.5144962	-0.2619071	0.0000000	-0.2091848	-0.4995760

شکل ۲. نتایج رویکرد مدل‌سازی پروبیت

ITEM1	ITEM2	ITEM3	ITEM4	ITEM5	ITEM6	ITEM7	ITEM8	ITEM9	ITEM10
-0.7459511	-0.6999351	-0.4475053	0.3035524	-0.8434880	-1.0519463	-0.5401135	0.0000000	-0.4361831	-1.0298752

شکل ۳. نتایج رویکرد مدل‌سازی لوجیت

از آنجایی که ضریب ITEM8 صفر درنظر گرفته شده است، سایر ضرایب تفاوت مقدار را با ضریب ITEM8 نشان می‌دهند. لذا از آنجایی که در هر دو مدل ضریب ITEM4 مثبت است در حالی‌که ضرایب بقیه ویژگی‌ها منفی هستند، این نتیجه بدست می‌آید که ITEM4 مهم‌تر از ITEM8 است و بقیه اهمیت کمتری نسبت به ITEM8 دارند.

probit

	shares
cost	0.09026214
Convenience	0.09096129
Safety against risks	0.10244832
Security and confidence	0.14706070
Honoring the customer	0.08500194
health care	0.07647632
To be fast	0.09845212
accessibility	0.12792925
flexibility	0.10378201
Dependence on technology	0.07762592
Sum	1.00000000

logit

	shares
cost	0.07492053
Convenience	0.07844862
Safety against risks	0.10097508
Security and confidence	0.21399044
Honoring the customer	0.06795808
health care	0.05517074
To be fast	0.09204388
accessibility	0.15796587
flexibility	0.10212483
Dependence on technology	0.05640195
Sum	1.00000000

شکل ٤. سهم ترجیحی ویژگی‌ها در دو مدل لوجیت و پروبیت

از سهم ترجیحی می‌توان به عنوان یک مقیاس برای تفسیر استفاده کرد. با توجه به شکل (٤) ترتیب ویژگی‌ها بر اساس رویکرد مدل‌سازی پروبیت و لوجیت و با استفاده از سهم ترجیحی با هم تفاوتی ندارند. ترتیب ویژگی‌ها در هر دو مدل به شرح جدول (۳) است.

جدول ۳. اولویت‌بندی ویژگی‌های تاکسی اینترنتی بر اساس مدل‌های پروبیت و لوجیت

اولویت	ویژگی	اولویت	ویژگی	اولویت	ویژگی
۱	امنیت و اطمینان خاطر	٥	سریع بودن	۹	وابستگی به فناوری (گوشی و اینترنت)
۲	سهولت دسترسی	٦	راحتی	۱۰	رعایت مسائل بهداشتی و فاصله‌گذاری اجتماعی
۳	انعطاف‌پذیری	۷	هزینه		
٤	ایمنی در مقابل خطرات	۸	تکریم مشتری		

همانطور که در شکل (٤) مشاهده می‌شود، در مدل لوجیت مهم‌ترین ویژگی امنیت و اطمینان خاطر (۰/۲۱۳۹۹۰٤٤) و دومین ویژگی سهولت دسترسی (۰/۱٥۷۹٦٥۸۷) به ترتیب تقریباً ۳/۹ و ۲/۹ برابر کمترین ویژگی یعنی رعایت مسائل بهداشتی و فاصله‌گذاری اجتماعی (۰/۰٥٥۱۷۰۷٤) اهمیت دارند. در مدل پروبیت نیز مهم‌ترین ویژگی امنیت و اطمینان خاطر (۰/۱٤۷۰٦۰۷۰) و دومین ویژگی سهولت دسترسی (۰/۱۲۷۹۲۹۲٥) به ترتیب تقریباً ۱/۹ و ۱/۷ برابر کمترین ویژگی یعنی رعایت مسائل بهداشتی و فاصله‌گذاری اجتماعی (۰/۰۷٦٤۷٦۳۲) اهمیت دارند.

ارزیابی مدل‌های پروبیت و لوجیت

شکل (۵) ماتریس درهم ریختگی در دو مدل پروبیت و لوجیت را نشان می‌دهد. همان‌طور که در این شکل مشخص است در فرآیند مدل‌سازی برای هر کدام از مدل‌های پروبیت و لوجیت جمعاً ۲۷۰۰۰ حالت (۲۷۰۰۰ زوج گزینه که هر زوج شامل بهترین و بدترین گزینه جهت انتخاب هر فرد می‌باشد) وجود دارد که ۱۱۸۱ حالت آن نمونه مثبت است. به عبارت دیگر نمونه مثبت حالتی می‌باشد که زوج مذکور که شامل بهترین و بدترین گزینه است توسط فرد انتخاب شده و ۲۵۸۱۹ حالت آن نمونه منفی است که بیانگر حالتی است که زوج گزینه موردنظر توسط فرد به عنوان بهترین و بدترین گزینه انتخاب نشده است. از میان ۱۱۸۱ نمونه مثبت تعداد نمونه‌های مثبتی که به درستی توسط مدل پروبیت به عنوان مثبت شناسایی شده‌اند ۹۲ نمونه و این تعداد توسط مدل لوجیت ۴۲۲ نمونه بوده است. تعداد نمونه‌های مثبتی که به اشتباه به عنوان منفی شناسایی شده‌اند، برای مدل پروبیت ۱۰۸۹ نمونه و برای مدل لوجیت ۷۵۹ نمونه می‌باشد. از میان ۲۵۸۱۹ نمونه منفی، تعداد نمونه‌های منفی که به درستی به عنوان منفی شناسایی شده‌اند برای مدل پروبیت ۲۴۷۱۱ نمونه و برای مدل لوجیت ۲۰۵۴۱ نمونه بوده است. همچنین تعداد نمونه‌های منفی که به اشتباه به‌عنوان مثبت شناسایی شده‌اند، برای مدل پروبیت ۱۱۰۸ نمونه و برای مدل لوجیت ۵۲۷۸ نمونه می‌باشد. با استفاده از مقادیر بدست آمده از شکل (۵) می‌توان معیار دقت را برای هر دو مدل پروبیت و لوجیت محاسبه کرد.

logit

	FALSE	TRUE
FALSE	20541	5278
TRUE	759	422

probit

	FALSE	TRUE
FALSE	24711	1108
TRUE	1089	92

شکل ۵. ماتریس در هم ریختگی در دو مدل لوجیت و پروبیت

جدول (٤) مقایسه معیارهای ارزیابی معرفی شده در این مطالعه را برای مدل‌های پروبیت و لوجیت نشان می‌دهد. همانطور که از داده‌های این جدول برمی‌آید معیار MSE نشانگر عملکرد بهتری برای مدل پروبیت است. در مورد معیار AIC نیز با تفاوت کمی مدل پروبیت بهتر عمل می‌کند. اما دقت مدل پروبیت ۹۱/۸ درصد است، در حالی‌که دقت مدل لوجیت ۷۷/٦ درصد است که از منظر این معیار نیز مدل پروبیت برتری دارد.

جدول ٤. مقایسه معیارهای ارزیابی در دو مدل لوجیت و پروبیت

نوع مدل	Accuracy	AIC	MSE
پروبیت	۰/۹۱۸٦۲۹٦	۱۱۱۸٦/٦۹	۰/۱۱٦٦٤٥۳
لوجیت	۰/۷۷٦٤۰۷٤	۱۱۲۰۱/۸۱	۰/۳۸۹٥۸۰۲

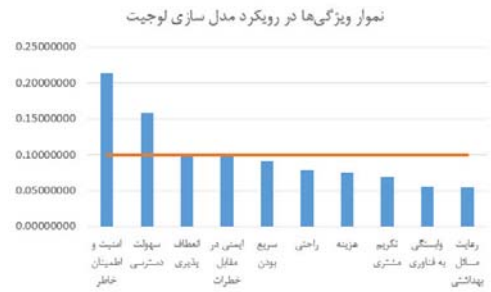
۵-نتیجه‌گیری

با توجه به نتایج می‌توان چنین نتیجه‌گیری کرد که اگرچه استفاده از مدل پروبیت برای مدل‌سازی عملکرد بهتری را نسبت به مدل لوجیت نشان می‌دهد، ولی نوع مدل تغییری را در رتبه‌بندی ویژگی‌های تاکسی اینترنتی ایجاد نمی‌نماید. اگر نمودار میله‌ای ویژگی‌های مختلف را بر اساس مقدار سهم ترجیحی ترسیم نماییم، موقیت هر ویژگی نسبت به خط

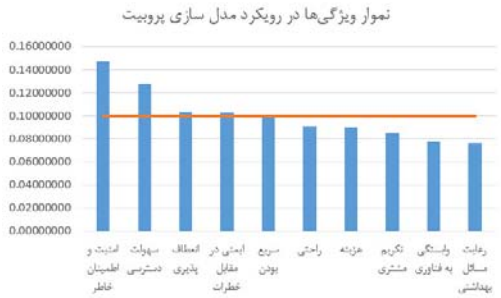
میانگین مشخص می‌شود. مشاهده می‌گردد ویژگی‌های امنیت و اطمینان خاطر، سهولت دسترسی، انعطاف‌پذیری (از نظر زمان و انتخاب مقصد واسطه و...) و ایمنی در مقابل خطرات بالاتر از خط میانگین و بقیه ویژگی‌ها پایین‌تر قرار دارند. در این راستا پیشنهاد می‌شود که به‌منظور بررسی دقیق‌تر مصداق‌های مرتبط با هر ویژگی، این مصداق‌ها در قالب سطوح مختلف برای این ویژگی‌ها تعریف شوند و با بهره‌گیری از م.ب.ب. نوع ۲ و ۳، امکان ارزیابی جامع‌تری در این زمینه فراهم گردد.

همچنین از شکل (٦) می‌توان نتیجه گرفت که ویژگی‌هایی از شیوه حمل‌ونقل تاکسی اینترنتی که بالاتر از خط میانگین قرار دارند، مانند امنیت و اطمینان خاطر و ایمنی در مقابل خطرات بیشتر به جوانب روحی و فکری مرتبط هستند و باعث آرامش روانی می‌شوند. همچنین ویژگی‌هایی مانند سهولت دسترسی و انعطاف‌پذیری (از نظر زمان و انتخاب مقصد واسطه و...) و سریع بودن نیز مطلوبیت‌هایی از شیوه حمل‌ونقل تاکسی اینترنتی را در مقایسه با سایر شیوه‌ها برجسته می‌سازند که موجبات آسودگی خاطر مسافران در بهره‌گیری از این شیوه را فراهم می نمایند که آنها نیز غیرمادی تلقی می‌شوند. در مقابل ویژگی‌هایی از قبیل هزینه، راحتی، تکریم مشتری، وابستگی به فناوری (گوشی و اینترنت) و رعایت مسائل بهداشتی و فاصله‌گذاری اجتماعی که پایین‌تر از خط میانگین قرار دارند، بیشتر به جنبه‌های مادی و اجتماعی و رفاه و سلامت جسمانی توجه دارند.

logit



probit



شکل ٦. نمودار میله‌ای ویژگی‌ها در رویکرد مدل‌سازی پروبیت و لوجیت و مقایسه با خط میانگین

٦-مراجع

-Ahmad, H., & Halim, H. (2017). Determining sample size for research activities. *Selangor Business Review*, 20-34.

-Aizaki, H., & Fogarty, J. (2023). R packages and tutorial for case 1 best–worst scaling. *Journal of Choice Modelling*, *46*, Article Number: 100394.

-Akbari, M., Amiri, N. S., Zúñiga, M. Á., Padash, H., & Shakiba, H. (2020). Evidence for Acceptance of Ride-Hailing Services in Iran. *Transportation Research Record*, *2674*(11), 289-303.

-Alemi, F., Circella, G., Handy, S., & Mokhtarian, P. (2018). What influences travelers to use Uber? Exploring the factors affecting the adoption of on-demand ride services in California [Article]. *Travel Behaviour and Society*, *13*, 88-104.

-Assele, S. Y., Meulders, M., & Vandebroek, M. (2023). Sample size selection for discrete choice experiments using design features. *Journal of Choice Modelling*, *49*, 100436.

-Burton, N., Burton, M., Rigby, D., Sutherland, C. A., & Rhodes, G. (2019). Best-worst scaling improves measurement of first impressions. *Cognitive research: principles and implications*, *4*(Article number: 36), 1-10.

-Chen, Y., Mahmassani, H. S., & Frei, A. (2018). Incorporating social media in travel

-Lancsar, E., Fiebig, D. G., & Hole, A. R. (2017). Discrete Choice Experiments: A Guide to Model Specification, Estimation and Software. *PharmacoEconomics*, *35*(7), 697-716.
-Li, H.-r. (2016). Taxi Positioning in the New Age of Internet and Industrial Development Research. *Procedia Engineering*, *137*, 811-816.
-Louviere, J., Lings, I., Islam, T., Gudergan, S., & Flynn, T. (2013). An introduction to the application of (case 1) best–worst scaling in marketing research. *International journal of research in marketing*, *30*(3), 292-303.
-Louviere, J. J., Flynn, T. N., & Marley, A. A. J. (2015). *Best-worst scaling: Theory, methods and applications*. Cambridge University Press.
-[Record #11 is using a reference type undefined in this output style.]
-Mjahed, L. B., Mittal, A., Elfar, A., Mahmassani, H. S., & Chen, Y. (2017). Exploring the role of social media platforms in informing trip planning: Case of Yelp.com [Article]. *Transportation Research Record*, *2666*, 1-9.
-Thompson, S. K. (1987). Sample Size for Estimating Multinomial Proportions. *The American Statistician*, *41*(1), 42-46.
-Walaza, S., Onderwater, P., & Zuidgeest, M. (2023). Best-worst scaling approach to measure public transport user quality perceptions and preferences in cape town.
-White, M. H. (2021). bwsTools: An R package for case 1 best-worst scaling. *Journal of Choice Modelling*, *39*, 100289.
-Yi, C., & Xia, L. (2015). A Forecasting model of trip distribution for vacant taxis with taxi-hailing Apps [J]. *Journal of Wuhan University of Technology (Transportation Science & Engineering)*, *39*(1), 51-54.
and activity choice models: conceptual framework and exploratory analysis. *International Journal of Urban Sciences*, *22*(2), 180-200.
-Choe, C., Jung, S., & Oaxaca, R. L. (2020). Identification and decompositions in probit and logit models. *Empirical Economics*, *59*, 1479-1492.
-Dow, C. R., Nguyen, D. B., Wang, S. C., Hwang, S. F., & Tsai, M. F. (2016). A Geo-Aware Taxi Carrying Management System by Using Location Based Services and Zone Queuing Techniques on Internet of Things [Article]. *Mobile Information Systems*, *2016*, Article ID 9817374, Article 9817374.
-Flynn, T. N., Louviere, J. J., Peters, T. J., & Coast, J. (2007). Best–worst scaling: What it can do for health care research and how to do it. *Journal of Health Economics*, *26*(1), 171-189.
-Guo, Q., & Shen, J. (2019). *An Empirical Comparison Between Discrete Choice Experiment and Best-worst Scaling: A Case Study of Mobile Payment Choice*. Research Institute for Economics and Business Administration, Kobe University.
-Hollis, G., & Westbury, C. (2018). When is best-worst best? A comparison of best-worst scaling, numeric estimation, and rating scales for collection of semantic norms. *Behavior Research Methods*, *50*, 115 - 133.
-Jing, P., Chen, Y., Wang, X., Pan, K., & Yuan, D. (2021). Evaluating the effectiveness of Didi ride-hailing security measures: An integration model. *Transportation Research Part F: Traffic Psychology and Behaviour*, *76*, 139-166.
-Koob, C. (2023). Consumers' Preferences for Digital Corporate Content on Company Websites: A Best–Worst Scaling Analysis. *Journal of Theoretical and Applied Electronic Commerce Research*, *18*(3), 1301-1319.

Comparison of Probit and Logit Modeling Approaches in Prioritizing Taxi Hailing Attributes: Best-Worst Scaling-Case 1

Mohsen Makaremi-Sharifi, Ph.D., Department of Transportation Planning, Faculty of Engineering, Imam Khomeini International University, Qazvin, Iran.
Amir Abbas Rassafi, Professor, Department of Transportation Planning, Faculty of Engineering, Imam Khomeini International University, Qazvin, Iran.

E-mail: m.makaremi.sh@gmail.com

Received: August 2024- Accepted: December 2024

ABSTRACT
The expansion of the internet and development in the business world have led to the emergence of new businesses, including taxi hailing. This service enables users to order taxis through a smartphone application, along with facilities such as online payment and driver evaluation. Competition in this field has driven taxi hailings to improve services and attract more customers, for example by understanding customer preferences and offering attributes that provide the best experience for them. The purpose of this study is to examine these attributes (10 attributes are selected here) and prioritize them. In this study, unlike the traditional approach of dealing with discrete choice models, which focuses on choosing the best (most important) alternative among alternatives, the role of the worst (least important) alternative is also considered. The probit model is employed for modeling, and its results are compared with the logit model.The results of this research show that although the coefficients of the features in the two models are not the same, and the performance evaluation criteria show the superiority of the probit model over the logit model, but the overall result in the rating of the features is not different.The results of this research indicate that although the coefficients of attributes in the two modeling approaches are not the same, and evaluation criteria demonstrate the superior performance of the probit model over the logit model, the overall conclusions regarding the ranking of attributes do not differ between them.When dealing with taxi hailing services, issues related to security and confidence are more important. In other words, users pay the most attention to aspects that bring about a sense of security and mental satisfaction when choosing taxi hailing services.Furthermore, this study indicates that issues such as compliance with health cares and social distancing are among the least important attributes.

Keywords: Probit, Logit, BWS-case1, Taxi Hailing Attributes