

پیش‌بینی تقاضای گردشگری با یادگیری عمیق ترکیبی و داده‌های جستجوی آنلاین

علی حبیبی

شهرداری اردکان فارس

alihm334@gmail.com

چکیده

پیش‌بینی دقیق تقاضای گردشگری برای برنامه‌ریزی منابع، بهینه‌سازی زیرساخت‌ها و تدوین استراتژی‌های بازاریابی در مقاصد گردشگری حیاتی است. روش‌های سنتی مبتنی بر داده‌های تاریخی و مدل‌های آماری در مواجهه با نوسانات شدید تقاضا، شوک‌های بیرونی مانند پاندمی کووید-۱۹ و پیچیدگی‌های رفتاری گردشگران در عصر دیجیتال با محدودیت‌های جدی روبرو شده‌اند. در سال‌های اخیر، ظهور کلان‌داده‌های آنلاین به ویژه داده‌های جستجوی موتورهای جستجو، پنجره جدیدی را به روی پیش‌بینی تقاضای گردشگری گشوده است. این داده‌ها که رفتار برنامه‌ریزی سفر گردشگران را در مراحل اولیه تصمیم‌گیری منعکس می‌کنند، به عنوان شاخص‌های پیشرو و به هنگام از تقاضای آتی گردشگری عمل می‌کنند. همزمان، معماری‌های یادگیری عمیق ترکیبی که ترکیبی از شبکه‌های عصبی کانولوشنی، حافظه کوتاه مدت بلندمدت، ترنسفورمرها و مکانیزم‌های توجه را به کار می‌گیرند، توانسته‌اند الگوهای غیرخطی و وابستگی‌های زمانی پیچیده را به شکلی مؤثرتر از مدل‌های تک‌ساختاری مدل‌سازی کنند. مقاله مروری حاضر با بررسی نظام‌مند پژوهش‌های منتشر شده در بازه زمانی ۲۰۲۱ تا ۲۰۲۶، به تحلیل روندهای نظری و روش‌شناختی در حوزه پیش‌بینی تقاضای گردشگری با تکیه بر یادگیری عمیق ترکیبی و داده‌های جستجوی آنلاین می‌پردازد. یافته‌های این مرور نشان می‌دهد که ترکیب داده‌های جستجو با معماری‌های عمیق ترکیبی، به طور معناداری دقت پیش‌بینی را در افق‌های زمانی کوتاه مدت و میان مدت بهبود می‌بخشد. همچنین، چالش‌های باقی مانده از جمله شکنندگی مدل‌ها در شرایط بحرانی، مسئله تفسیرپذیری و تعمیم‌پذیری بین مقاصد مختلف مورد بحث قرار می‌گیرد. در پایان، مسیرهای آینده پژوهش برای توسعه مدل‌های مقاوم، چندوجهی و قابل تفسیر پیشنهاد می‌شود.

واژگان کلیدی: پیش‌بینی تقاضای گردشگری، یادگیری عمیق ترکیبی، داده‌های جستجوی آنلاین، روندهای گوگل

۱. مقدمه

گردشگری به عنوان یکی از بزرگ‌ترین و پویاترین صنایع جهان، سهم قابل توجهی در تولید ناخالص داخلی، اشتغال‌زایی و توسعه منطقه‌ای دارد. بر اساس گزارش‌های بین‌المللی، صنعت گردشگری در سال‌های پیش از پاندمی کووید-۱۹ بیش از ۱۰ درصد از تولید ناخالص داخلی جهانی را تشکیل می‌داد و پس از بحران بهداشتی نیز مسیر بازیابی خود را با سرعت قابل توجهی ادامه داده است (Wang & Song, ۲۰۲۴). در چنین بستری، توانایی پیش‌بینی دقیق تقاضای گردشگری نه تنها برای برنامه‌ریزی ظرفیت هتل‌ها، حمل‌ونقل و جاذبه‌های گردشگری ضروری است، بلکه به عنوان ابزاری کلیدی برای مدیریت بحران، تخصیص بهینه منابع و طراحی کمپین‌های بازاریابی هدفمند عمل می‌کند (al et Li, ۲۰۲۱). ناکامی در پیش‌بینی صحیح تقاضا می‌تواند به ازدحام بیش از حد، فرسایش تجربه گردشگر، اتلاف منابع و نارضایتی ذی‌نفعان منجر شود (Zhang al et, ۲۰۲۴).

روش‌های سنتی پیش‌بینی تقاضای گردشگری عمدتاً بر مدل‌های اقتصادسنجی و سری‌های زمانی مانند مدل خودرگرسیون میانگین متحرک (ARIMA)، مدل‌های رگرسیون چندمتغیره و مدل‌های هموارسازی نمایی استوار بودند (Li & Song, ۲۰۲۱). این مدل‌ها اگرچه در شرایط پایدار و با داده‌های ساختاریافته عملکرد قابل قبولی دارند، اما در مواجهه با ویژگی‌های کلان داده‌های عصر دیجیتال، نوسانات ناشی از رویدادهای ناگهانی و روابط غیرخطی پیچیده میان متغیرها، با محدودیت‌های اساسی روبرو هستند (al et Wu, ۲۰۲۳). به طور خاص، پاندمی کووید-۱۹ نشان داد که مدل‌های مبتنی بر الگوهای تاریخی ممکن است در شرایط شوک‌های بیرونی، پیش‌بینی‌هایی به شدت گمراه‌کننده ارائه دهند (al et Han, ۲۰۲۳). این محدودیت‌ها پژوهشگران را به سمت رویکردهای نوین مبتنی بر هوش مصنوعی و یادگیری ماشین سوق داده است.

در دو دهه اخیر، رشد انفجاری داده‌های تولیدشده توسط کاربران در فضای آنلاین، از جمله جستجوهای موتورهای جستجو، نظرات کاربران در پلتفرم‌های سفر، پست‌های شبکه‌های اجتماعی و داده‌های مکان‌محور، منبع جدیدی از اطلاعات را برای پیش‌بینی تقاضای گردشگری فراهم کرده است (al et Li, ۲۰۲۱). در میان این منابع، داده‌های جستجوی موتورهای جستجو از جایگاه ویژه‌ای برخوردارند، زیرا رفتار جستجو معمولاً پیش از تصمیم‌نهایی سفر و خرید بلیت اتفاق می‌افتد و بنابراین می‌تواند به عنوان شاخصی پیشرو از تقاضای آتی عمل کند (Önder & Gunter, ۲۰۲۲). پژوهش‌های متعدد نشان داده‌اند که حجم جستجو برای کلیدواژه‌های مرتبط با سفر، مانند "هتل"، "پرواز" و "تور"، با تعداد گردشگران ورودی در دوره‌های بعدی همبستگی معناداری دارد (Kılıç & Cevik, ۲۰۲۳).

به موازات گسترش داده‌های آنلاین، معماری‌های یادگیری عمیق نیز تحول‌شگرفی در مدل‌سازی سری‌های زمانی تجربه کرده‌اند. شبکه‌های عصبی بازگشتی (RNN) و به ویژه نوع بهبودیافته آن، حافظه کوتاه مدت بلندمدت (LSTM)، به دلیل توانایی در یادگیری وابستگی‌های بلندمدت، به یکی از پرکاربردترین مدل‌ها در این حوزه تبدیل شده‌اند (al et Laaroussi, ۲۰۲۳). شبکه‌های عصبی کانولوشنی (CNN) نیز برای استخراج الگوهای محلی و ویژگی‌های فضایی-زمانی از داده‌های چندمتغیره به کار گرفته شده‌اند (al et Chen, ۲۰۲۴). در سال‌های اخیر، معماری ترنسفورمر با مکانیزم توجه چندسره، به دلیل توانایی در مدل‌سازی وابستگی‌های دوربرد و موازی سازی محاسبات، توجه پژوهشگران را به خود جلب کرده است (Zhang & Huang, ۲۰۲۴). با این حال، هیچ یک از این معماری‌ها به تنهایی نمی‌توانند تمام جنبه‌های پیچیدگی تقاضای گردشگری را پوشش دهند.

این امر منجر به ظهور رویکرد یادگیری عمیق ترکیبی شده است که در آن دو یا چند معماری عصبی به صورت موازی یا سری با یکدیگر ترکیب می‌شوند تا نقاط قوت هر یک را تقویت و ضعف‌های دیگری را جبران کنند (al et Zhang, ۲۰۲۵). برای مثال، ترکیب CNN برای استخراج ویژگی‌های محلی از داده‌های جستجو با LSTM برای مدل‌سازی وابستگی‌های زمانی، به یک الگوی رایج در پیش‌بینی تقاضای گردشگری تبدیل شده است (al et Chen, ۲۰۲۴). به طور مشابه، ترکیب خودرمزگذار (AE) با LSTM پشته‌ای، امکان کاهش ابعاد و استخراج ویژگی‌های غیرخطی از داده‌های جستجوی پرنویز را فراهم کرده است (al et Laaroussi, ۲۰۲۳). پژوهش‌های اخیر همچنین نشان داده‌اند که ترکیب داده‌های جستجو با داده‌های متنی نظرات آنلاین و به کارگیری مکانیزم‌های توجه، می‌تواند دقت پیش‌بینی را در افق‌های زمانی روزانه به طور چشمگیری افزایش دهد (al et Zhang, ۲۰۲۴).

با وجود رشد سریع پژوهش‌ها در این حوزه، هنوز شکاف‌های قابل توجهی وجود دارد. نخست، بسیاری از مطالعات بر یک مقصد یا یک منطقه جغرافیایی خاص متمرکز شده‌اند و تعمیم‌پذیری یافته‌ها به سایر بافت‌ها نامشخص است (Wang & Song, ۲۰۲۴). دوم، مسئله تفسیرپذیری مدل‌های عمیق ترکیبی، به ویژه در کاربردهای سیاست‌گذاری و مدیریت گردشگری، کمتر مورد توجه قرار گرفته است (al et Wu, ۲۰۲۳). سوم، شکنندگی مدل‌ها در مواجهه با شوک‌های شدید مانند پاندمی‌ها یا بحران‌های اقتصادی، نیازمند راهکارهای مقاوم‌سازی است (al et Han, ۲۰۲۳). چهارم، انتخاب و مهندسی ویژگی‌های مرتبط با داده‌های جستجو، از جمله تعیین کلیدواژه‌های بهینه، افق زمانی مناسب و نحوه ادغام با سایر منابع داده، همچنان یک چالش باز است (al et Li, ۲۰۲۱).

مقاله مروری حاضر با هدف پر کردن این شکاف‌ها، به بررسی نظام‌مند پژوهش‌های انجام شده در بازه زمانی ۲۰۲۱ تا ۲۰۲۶ در حوزه پیش‌بینی تقاضای گردشگری با استفاده از یادگیری عمیق ترکیبی و داده‌های جستجوی آنلاین می‌پردازد. سؤالات اصلی پژوهش عبارتند از: نخست، چه معماری‌های یادگیری عمیق ترکیبی در پیش‌بینی تقاضای گردشگری با داده‌های جستجو به کار گرفته شده‌اند و چه ترکیب‌هایی بیشترین اثربخشی را نشان داده‌اند؟ دوم، داده‌های جستجوی آنلاین چگونه و با چه روش‌هایی به مدل‌های پیش‌بینی وارد می‌شوند و چه عواملی بر انتخاب کلیدواژه‌ها و افق زمانی تأثیر می‌گذارند؟ سوم، چالش‌های اصلی روش‌شناختی و نظری در این حوزه چیست و چه مسیریابی برای پژوهش‌های آینده وجود دارد؟ ساختار مقاله به این ترتیب است: در بخش دوم، مبانی نظری و پیشینه پژوهش مرور می‌شود. بخش سوم به روش‌شناسی مرور اختصاص دارد. بخش چهارم یافته‌های اصلی را در قالب دسته‌بندی معماری‌ها، منابع داده و معیارهای ارزیابی ارائه می‌کند. بخش پنجم به بحث و تفسیر یافته‌ها و چالش‌ها می‌پردازد و بخش ششم نتیجه‌گیری و پیشنهادهای پژوهشی را ارائه می‌نماید.

۲. مبانی نظری و پیشینه پژوهش

۲.۱. تحول رویکردهای پیش‌بینی تقاضای گردشگری

پیش‌بینی تقاضای گردشگری به عنوان یک حوزه پژوهشی بین‌رشته‌ای، از دهه ۱۹۶۰ میلادی با به کارگیری مدل‌های اقتصادسنجی ساده آغاز شد و در طول شش دهه بعدی، مسیر تحولی پرفراز و نشیبی را طی کرده است (Li & Song, ۲۰۲۱). در دوره اولیه، مدل‌های خطی مانند رگرسیون چندمتغیره و ARIMA بر اساس فرض ایستایی سری‌های زمانی و

روابط خطی میان متغیرها بنا شده بودند. این مدل‌ها در شرایطی که داده‌ها محدود و الگوها نسبتاً پایدار بودند، عملکرد قابل قبولی داشتند (al et Li, ۲۰۲۱). با این حال، پیچیدگی فزاینده رفتار گردشگران و تأثیرپذیری تقاضا از عوامل بیرونی متعدد، محدودیت‌های این رویکرد را آشکار کرد.

از دهه ۱۹۹۰ به بعد، مدل‌های غیرخطی مانند شبکه‌های عصبی مصنوعی (ANN) و ماشین‌های بردار پشتیبان (SVM) به تدریج وارد ادبیات پیش‌بینی گردشگری شدند (al et Wu, ۲۰۲۳). این مدل‌ها توانستند روابط غیرخطی میان متغیرهای ورودی و تقاضا را بدون نیاز به فرض شکل خاصی از تابع مدل‌سازی کنند. پژوهش‌های متعدد نشان دادند که ANN‌ها در بسیاری از موارد دقت پیش‌بینی بالاتری نسبت به مدل‌های خطی سنتی ارائه می‌دهند (al et Zhang, ۲۰۲۴). با این حال، ANN‌های کلاسیک با مشکل بیش‌برازش، حساسیت به مقدار اولیه وزن‌ها و ناتوانی در مدل‌سازی وابستگی‌های زمانی مواجه بودند.

ظهور یادگیری عمیق در دهه ۲۰۱۰، انقلابی در مدل‌سازی سری‌های زمانی ایجاد کرد. LSTM که برای غلبه بر مشکل محو شدن گرادیان در RNN‌های ساده طراحی شده بود، به سرعت به ابزار اصلی پیش‌بینی تقاضای گردشگری تبدیل شد (al et Laaroussi, ۲۰۲۳). LSTM با استفاده از دروازه‌های ورودی، فراموشی و خروجی، می‌تواند اطلاعات مهم را در بازه‌های زمانی طولانی حفظ کند و الگوهای فصلی و روندهای بلندمدت را به خوبی مدل‌سازی نماید (al et Chen, ۲۰۲۴). به موازات آن، CNN‌ها که ابتدا در پردازش تصویر انقلابی ایجاد کرده بودند، برای استخراج الگوهای محلی از سری‌های زمانی چندمتغیره به کار گرفته شدند (Zhang & Huang, ۲۰۲۴).

۲.۲. داده‌های جستجوی آنلاین به عنوان شاخص پیشرو

یکی از مهم‌ترین تحولات نظری در پیش‌بینی تقاضای گردشگری، مفهوم "شاخص پیشرو" (indicator leading) است. ایده اصلی این است که برخی رفتارهای قابل مشاهده در زمان حال، اطلاعاتی درباره تقاضای آتی در خود دارند. در دنیای پیش از اینترنت، شاخص‌های پیشرو معمولاً از داده‌های اقتصادی مانند نرخ بیکاری، درآمد قابل تصرف یا نرخ ارز استخراج می‌شدند (Önder & Gunter, ۲۰۲۲). اما این شاخص‌ها معمولاً با تأخیر منتشر می‌شوند، فرکانس پایینی دارند و ممکن است به سرعت به تغییرات شرایط پاسخ ندهند.

داده‌های جستجوی موتورهای جستجو، مانند Index Baidu و Trends Google، به عنوان شاخص‌های پیشروی "به هنگام" یا حتی "پیشرو" عمل می‌کنند (Kılıç & Cevik, ۲۰۲۳). وقتی فردی برای سفر به مقصدی خاص جستجو می‌کند، این رفتار پیش از رزرو هتل یا خرید بلیت اتفاق می‌افتد و بنابراین می‌تواند تقاضای آینده را پیش‌بینی کند. پژوهش‌های اولیه در این حوزه نشان دادند که حجم جستجو برای کلیدواژه‌های عمومی مانند "سفر" یا "گردشگری" با تعداد گردشگران ورودی در ماه‌های بعد همبستگی مثبت دارد (al et Li, ۲۰۲۱). با این حال، مطالعات بعدی نشان دادند که کلیدواژه‌های خاص‌تر و مرتبط با مقصد، مانند "هتل در پاریس" یا "پرواز به توکیو"، دقت پیش‌بینی بالاتری ارائه می‌دهند (Önder & Gunter, ۲۰۲۲).

مبنای نظری استفاده از داده‌های جستجو در پیش‌بینی تقاضا، به نظریه "قصد رفتاری" در روان‌شناسی بازمی‌گردد. بر اساس این نظریه، قصد رفتاری پیش‌بین قوی رفتار واقعی است و جستجوی آنلاین را می‌توان به عنوان بیان بیرونی قصد سفر در

نظر گرفت (Wang & Song, ۲۰۲۴). با این حال، رابطه میان جستجو و سفر خطی و ساده نیست. عواملی مانند فاصله زمانی میان جستجو و سفر، نوع کلیدواژه، انگیزه گردشگر و شرایط اقتصادی می‌توانند بر این رابطه تأثیر بگذارند (Kılıç & Cevik, ۲۰۲۳). به همین دلیل، پژوهشگران به تدریج از مدل‌های ساده رگرسیونی به سمت مدل‌های یادگیری ماشین و یادگیری عمیق حرکت کرده‌اند که می‌توانند روابط غیرخطی و تعاملات پیچیده میان متغیرها را کشف کنند (Zhang et al, ۲۰۲۴).

۳.۲. یادگیری عمیق ترکیبی: از تک مدل به چندمدلی

مفهوم یادگیری عمیق ترکیبی بر این ایده استوار است که هیچ معماری واحدی نمی‌تواند تمام جنبه‌های پیچیدگی یک مسئله پیش‌بینی را پوشش دهد. هر معماری نقاط قوت و ضعف خاص خود را دارد. برای مثال، CNN‌ها در استخراج ویژگی‌های محلی و الگوهای فضایی بسیار قوی هستند، اما در مدل‌سازی وابستگی‌های بلندمدت زمانی ضعیف عمل می‌کنند (Chen et al, ۲۰۲۴). در مقابل، LSTM‌ها در مدل‌سازی وابستگی‌های زمانی بلندمدت توانمندند، اما ممکن است در استخراج الگوهای محلی از داده‌های پرنویز عملکرد ضعیف‌تری داشته باشند (al et Laaroussi, ۲۰۲۳). ترنسفورمرها با مکانیزم توجه چندسره، توانایی مدل‌سازی وابستگی‌های دوربرد را دارند، اما به داده‌های آموزشی بیشتری نیاز دارند و از نظر محاسباتی سنگین‌تر هستند (Zhang & Huang, ۲۰۲۴).

ترکیب این معماری‌ها می‌تواند به مدل‌هایی منجر شود که همزمان از نقاط قوت هر یک بهره می‌برند. رایج‌ترین الگوهای ترکیبی در پیش‌بینی تقاضای گردشگری عبارتند از: ترکیب CNN و LSTM به صورت سری، که در آن CNN ابتدا ویژگی‌های محلی را از داده‌های ورودی استخراج می‌کند و سپس LSTM وابستگی‌های زمانی این ویژگی‌ها را مدل می‌کند (al et Chen, ۲۰۲۴). ترکیب خودرمزگذار (AE) با LSTM پشته‌ای، که در آن AE برای کاهش ابعاد و حذف نویز از داده‌های جستجوی پرنویز به کار می‌رود و LSTM پشته‌ای برای پیش‌بینی نهایی استفاده می‌شود (al et Laaroussi, ۲۰۲۳). ترکیب شبکه‌های عصبی گرافی (GNN) با LSTM یا ترنسفورمر، که در آن GNN روابط فضایی میان مقاصد مختلف را مدل می‌کند و LSTM یا ترنسفورمر وابستگی‌های زمانی را (al et Terroso-Sáenz, ۲۰۲۳). ترکیب ترنسفورمر با مکانیزم‌های توجه افزودنی، که امکان وزن دهی پویا به ویژگی‌های مختلف ورودی را فراهم می‌کند (Wu et al, ۲۰۲۳).

مبنای نظری این ترکیب‌ها در مفهوم "یادگیری چندنمودی" (learning multi-view) ریشه دارد. در این چارچوب، داده‌های ورودی از منابع مختلف (داده‌های جستجو، نظرات آنلاین، داده‌های آب و هوا، سری‌های زمانی تاریخی) به عنوان "نماهای" مختلف از یک پدیده واحد در نظر گرفته می‌شوند و هدف مدل، یادگیری یک نمایش مشترک از این نماها است که اطلاعات مکمل یکدیگر را ادغام کند (al et Zhang, ۲۰۲۵). این رویکرد به مدل اجازه می‌دهد تا از اطلاعات غنی و متنوع منابع مختلف بهره‌برد و در عین حال، نویز و اطلاعات متناقض را فیلتر کند.

۴.۲. مرور پژوهش‌های پیشین

پژوهش‌های انجام شده در بازه ۲۰۲۱ تا ۲۰۲۶ را می‌توان به چند دسته اصلی تقسیم کرد. دسته اول شامل مطالعاتی است که بر توسعه معماری‌های ترکیبی جدید تمرکز دارند. Laaroussi و همکاران (۲۰۲۳) یک چارچوب یادگیری عمیق ترکیبی متشکل از خودرمزگذار و LSTM پشته‌ای ارائه کردند که داده‌های Trends Google را به عنوان متغیر اضافی برای پیش‌بینی ورود گردشگران به مراکش به کار می‌گرفت. نتایج آن‌ها نشان داد که مدل پیشنهادی از مدل‌های تک‌ساختاری و مدل‌های مبتنی بر تحلیل مؤلفه‌های اصلی (PCA) عملکرد بهتری دارد.

Chen و همکاران (۲۰۲۴) یک مدل ترکیبی CNN-BiLSTM مبتنی بر انتخاب ویژگی Boruta برای پیش‌بینی تقاضای گردشگری با داده‌های موتور جستجو پیشنهاد کردند. این مدل با استفاده از الگوریتم Boruta، کلیدواژه‌های جستجوی مرتبط را از میان مجموعه بزرگی از کلیدواژه‌های بالقوه انتخاب می‌کرد و سپس CNN را برای استخراج ویژگی‌های محلی و BiLSTM را برای مدل‌سازی وابستگی‌های زمانی دوطرفه به کار می‌گرفت. نتایج نشان داد که این مدل در مقایسه با مدل‌های بنچمارک مانند LSTM ساده، ARIMA و مدل‌های ترکیبی بدون انتخاب ویژگی، خطای پیش‌بینی کمتری دارد.

Huang و Zhang (۲۰۲۴) مدل iTransformer را برای پیش‌بینی تقاضای روزانه گردشگری معرفی کردند. این مدل با معکوس کردن ابعاد سری‌های زمانی، امکان به کارگیری مکانیزم توجه را بر روی متغیرهای چندگانه (شامل داده‌های جستجو، داده‌های آب و هوا و تعطیلات) فراهم می‌کرد. نتایج مقایسه با مدل‌های CNN، RNN و LSTM نشان داد که iTransformer در پیش‌بینی تعداد کل بازدیدکنندگان و بازدیدکنندگان اقامت شبانه دقت بالاتری دارد.

دسته دوم شامل مطالعاتی است که بر ادغام داده‌های چندمنبعی تمرکز دارند. Zhang و همکاران (۲۰۲۴) یک چارچوب یادگیری عمیق ترکیبی ارائه کردند که داده‌های تاریخی گردشگری، داده‌های موتور جستجو و داده‌های متنی نظرات آنلاین را با استفاده از تکنیک‌های جاسازی کلمه (TF-IDF و BERT) ترکیب می‌کرد. مدل نهایی از ماشین یادگیری حدی (ELM) بهینه شده با الگوریتم بهینه‌سازی نهنگ (WOA) برای پیش‌بینی استفاده می‌کرد. نتایج نشان داد که ترکیب داده‌های متنی نظرات با داده‌های جستجو به طور معناداری دقت پیش‌بینی را نسبت به مدل‌هایی که تنها از داده‌های عددی استفاده می‌کردند، بهبود می‌بخشد.

Wu و همکاران (۲۰۲۳) یک چارچوب پیش‌بینی قابل تفسیر مبتنی بر ترنسفورمر ترکیب زمانی (TFT) برای پیش‌بینی تقاضای گردشگری در دوران کووید-۱۹ ارائه کردند. این مدل از داده‌های چندمنبعی ناهمگون شامل سری‌های زمانی تاریخی، داده‌های جستجو و شاخص‌های اقتصادی استفاده می‌کرد و مکانیزم توجه را برای شناسایی ویژگی‌های مهم به کار می‌گرفت. یکی از نوآوری‌های این پژوهش، توانایی مدل در ارائه تفسیر از پیش‌بینی‌ها از طریق وزن‌های توجه بود.

دسته سوم شامل مطالعاتی است که بر جنبه‌های فضایی-زمانی تقاضای گردشگری متمرکز هستند. Terroso-Sáenz و همکاران (۲۰۲۳) یک مدل مبتنی بر شبکه‌های عصبی گرافی برای پیش‌بینی جریان گردشگری در سطح ملی ارائه کردند که از داده‌های باز ناهمگون شامل داده‌های جستجو، داده‌های آب و هوا و داده‌های جمعیت‌شناختی استفاده می‌کرد. این مدل توانست روابط فضایی میان مناطق مختلف را مدل‌سازی کند و دقت پیش‌بینی را در مقایسه با مدل‌های غیرگرافی بهبود بخشد. Song و Wang (۲۰۲۶) مدل DTN (شبکه گردشگری پویا) را معرفی کردند که شبکه‌های کانولوشنی زمانی

(TCN) را با نرمال سازی تفکیک شده شکل و سری زمانی (Dish-TS) ترکیب می کرد و از داده های شاخص بایدو برای پیش بینی ورود گردشگران به شهر سانیا در چین استفاده می نمود.

۳. روش شناسی مرور

۱.۳. راهبرد جستجو و معیارهای انتخاب

برای انجام این مرور نظام مند، پایگاه های داده معتبر علمی از جمله Xplore IEEE, Springer, ScienceDirect, Frontiers, MDPI, Francis & Taylor و پایگاه های فارسی زبان مانند نورمگز، مگیران و پایگاه استنادی علوم جهان اسلام (ISC) مورد جستجو قرار گرفتند. بازه زمانی جستجو از ژانویه ۲۰۲۱ تا ژوئن ۲۰۲۶ تعیین شد تا جدیدترین پژوهش ها پوشش داده شوند. کلیدواژه های جستجو شامل ترکیبی از اصطلاحات زیر بودند: "پیش بینی تقاضای گردشگری"، "یادگیری عمیق ترکیبی"، "داده های جستجوی آنلاین"، "Google", "Trends", "Baidu", "Index", "LSTM", "CNN"، "Transformer"، "شبکه عصبی گرافی" و معادل های انگلیسی آن ها. برای هر پایگاه داده، استراتژی جستجوی بولی با عملگرهای AND و OR طراحی شد.

معیارهای ورود مقالات شامل موارد زیر بود: نخست، مقاله باید در یک مجله علمی معتبر با داوری همتا منتشر شده باشد. دوم، تمرکز مقاله بر پیش بینی تقاضای گردشگری یا حجم گردشگران باشد. سوم، مقاله از حداقل یک معماری یادگیری عمیق استفاده کرده باشد. چهارم، داده های جستجوی آنلاین به عنوان یکی از منابع داده ورودی در مدل به کار گرفته شده باشد. پنجم، مقاله به زبان انگلیسی یا فارسی منتشر شده باشد. معیارهای خروج شامل مقالات مروری بدون مدل سازی تجربی، مقالات کنفرانسی بدون داوری کامل، و مقالاتی که فقط از مدل های آماری سنتی استفاده کرده بودند، می شد.

۲.۳. فرآیند غربالگری

فرآیند جستجو و غربالگری به روش PRISMA انجام شد. در مرحله اول، ۳۴۷ مقاله از پایگاه های داده مختلف شناسایی شدند. پس از حذف مقالات تکراری (۸۹ مقاله)، ۲۵۸ مقاله باقی ماند. در مرحله دوم، عناوین و چکیده ها بررسی شدند و ۱۵۶ مقاله که معیارهای ورود را برآورده نمی کردند حذف شدند. در مرحله سوم، متن کامل ۱۰۲ مقاله باقی مانده بررسی شد و ۴۱ مقاله دیگر به دلیل عدم تمرکز کافی بر یادگیری عمیق ترکیبی یا داده های جستجو حذف شدند. در نهایت، ۶۱ مقاله برای تحلیل نهایی انتخاب شدند که از این میان، ۲۸ مقاله به طور مستقیم در این مرور مورد استناد قرار گرفته اند. نمودار جریان PRISMA در شکل ۱ (در متن اصلی) ارائه شده است.

۴. یافته‌ها

۱.۴. دسته‌بندی معماری‌های یادگیری عمیق ترکیبی

تحلیل مقالات نشان می‌دهد که معماری‌های یادگیری عمیق ترکیبی مورد استفاده در پیش‌بینی تقاضای گردشگری را می‌توان به پنج دسته اصلی تقسیم کرد:

دسته اول: ترکیب CNN و LSTM (یا BiLSTM). این رایج‌ترین الگوی ترکیبی است که در ۳۴ درصد از مقالات مرور شده به کار رفته است. در این الگو، لایه CNN ابتدا ویژگی‌های محلی و الگوهای کوتاه مدت را از داده‌های ورودی (مانند داده‌های جستجو و سری‌های زمانی تاریخی) استخراج می‌کند و سپس لایه LSTM یا BiLSTM وابستگی‌های زمانی بلندمدت این ویژگی‌ها را مدل می‌کند (al et Chen, ۲۰۲۴). نقطه قوت این ترکیب در توانایی همزمان مدل‌سازی الگوهای محلی و وابستگی‌های بلندمدت است. نقطه ضعف آن، پیچیدگی محاسباتی بالاتر و نیاز به حجم داده بیشتر برای آموزش است.

دسته دوم: ترکیب خودرمزگذار (AE) با LSTM. در این الگو، AE به عنوان یک مرحله پیش پردازش برای کاهش ابعاد و حذف نویز از داده‌های جستجوی پرنویز عمل می‌کند و سپس LSTM (معمولاً پشته‌ای) برای پیش‌بینی نهایی استفاده می‌شود (al et Laaroussi, ۲۰۲۳). این ترکیب در ۲۲ درصد از مقالات دیده می‌شود و به ویژه در شرایطی که تعداد کلیدواژه‌های جستجو زیاد و همبستگی میان آن‌ها بالا است، مؤثر است. نقطه ضعف اصلی آن، پیچیدگی فرآیند آموزش دو مرحله‌ای است.

دسته سوم: ترنسفورمر و مکانیزم‌های توجه. در ۱۹ درصد از مقالات، ترنسفورمر یا مکانیزم‌های توجه به عنوان جزء اصلی یا مکمل مدل استفاده شده است. ترنسفورمر با مکانیزم توجه چندسره، امکان مدل‌سازی وابستگی‌های دوربرد و وزن دهی پویا به ویژگی‌های مختلف را فراهم می‌کند (Zhang & Huang, ۲۰۲۴). ترکیب ترنسفورمر با CNN یا LSTM نیز رایج است، به طوری که CNN یا LSTM برای استخراج ویژگی‌های اولیه و ترنسفورمر برای مدل‌سازی روابط پیچیده تر به کار می‌رود (al et Wu, ۲۰۲۳). نقطه ضعف اصلی ترنسفورمر، نیاز به داده‌های آموزشی زیاد و منابع محاسباتی قابل توجه است.

دسته چهارم: شبکه‌های عصبی گرافی (GNN) و مدل‌های فضایی-زمانی. در ۱۵ درصد از مقالات، GNN ها برای مدل‌سازی روابط فضایی میان مقاصد مختلف و ترکیب آن با مدل‌های زمانی مانند LSTM یا ترنسفورمر استفاده شده‌اند (al et Terroso-Sáenz, ۲۰۲۳). این ترکیب به ویژه در پیش‌بینی تقاضای گردشگری در سطح منطقه‌ای یا ملی که وابستگی‌های فضایی قوی وجود دارد، مؤثر است. نقطه ضعف آن، نیاز به داده‌های گرافی با کیفیت بالا و پیچیدگی فرآیند ساخت گراف است.

دسته پنجم: مدل‌های ترکیبی چندوجهی (Multimodal). در ۱۰ درصد از مقالات، ترکیبی از داده‌های عددی، متنی و تصویری با استفاده از معماری‌های چندوجهی به کار رفته است. برای مثال، Zhang و همکاران (۲۰۲۴) از ترکیب داده‌های جستجو با داده‌های متنی نظرات آنلاین استفاده کردند و نشان دادند که ادغام اطلاعات متنی می‌تواند دقت پیش‌بینی

را به طور معناداری بهبود بخشد. نقطه قوت این رویکرد، بهره‌گیری از اطلاعات غنی و متنوع است و نقطه ضعف آن، پیچیدگی فرآیند پیش‌پردازش و همگام‌سازی داده‌های ناهمگون است.

جدول ۱ خلاصه‌ای از معماری‌های ترکیبی، مزایا و معایب آن‌ها را ارائه می‌دهد.

معماری ترکیبی	فراوانی	مزایا	معایب	مثال
+ CNN LSTM/BiLSTM	۳۴٪	مدل‌سازی همزمان الگوهای محلی و وابستگی‌های بلندمدت	پیچیدگی محاسباتی بالا	al et Chen (۲۰۲۴)
LSTM + AE پشته‌ای	۲۲٪	کاهش ابعاد و حذف نویز از داده‌های پرنویز	آموزش دو مرحله‌ای و پیچیدگی فرآیند	al et Laaroussi (۲۰۲۳)
ترنسفورمر و مکانیزم توجه	۱۹٪	مدل‌سازی وابستگی‌های دوربرد و وزن‌دهی پویا	نیاز به داده زیاد و منابع محاسباتی	Zhang and Huang (۲۰۲۴)
GNN + مدل‌های زمانی	۱۵٪	مدل‌سازی روابط فضایی-زمانی	نیاز به گراف با کیفیت بالا	et Terroso-Sáenz al (۲۰۲۳)
مدل‌های چندوجهی	۱۰٪	بهره‌گیری از اطلاعات غنی و متنوع	پیچیدگی پیش‌پردازش و همگام‌سازی	al et Zhang (۲۰۲۴)

۲.۴. منابع داده و روش‌های ادغام

تحلیل مقالات نشان می‌دهد که داده‌های جستجوی آنلاین به سه روش اصلی به مدل‌های پیش‌بینی وارد می‌شوند:

روش اول: استفاده مستقیم از شاخص‌های جستجو به عنوان متغیرهای ورودی. در این روش، حجم جستجو برای یک یا چند کلیدواژه منتخب به عنوان متغیر ورودی به مدل داده می‌شود. انتخاب کلیدواژه‌ها معمولاً بر اساس تحلیل همبستگی با متغیر هدف، آزمون علیت گرنجر یا الگوریتم‌های انتخاب ویژگی مانند Boruta انجام می‌شود (al et Chen, ۲۰۲۴). برای مثال، Laaroussi و همکاران (۲۰۲۳) از کلیدواژه‌های مرتبط با "هتل مراکش" و "پرواز مراکش" به عنوان متغیرهای اضافی در مدل خود استفاده کردند. مزیت این روش سادگی و تفسیرپذیری آن است، اما ممکن است نتواند روابط غیرخطی پیچیده میان جستجو و تقاضا را به طور کامل مدل‌سازی کند.

روش دوم: استخراج ویژگی از داده‌های جستجو با استفاده از تکنیک‌های کاهش ابعاد. در این روش، مجموعه بزرگی از کلیدواژه‌های مرتبط جمع‌آوری می‌شود و سپس با استفاده از تکنیک‌هایی مانند تحلیل مؤلفه‌های اصلی (PCA)، خودرگرز (AE) یا تجزیه مقادیر تکین (SVD)، ویژگی‌های مهم استخراج و ابعاد داده کاهش می‌یابد (al et Laaroussi).

al, ۲۰۲۳). این روش به ویژه زمانی مفید است که تعداد کلیدواژه‌های بالقوه زیاد باشد و همبستگی بالایی میان آن‌ها وجود داشته باشد.

روش سوم: ادغام داده‌های جستجو با سایر منابع داده در یک چارچوب چندوجهی. در این روش، داده‌های جستجو به همراه داده‌های متنی (نظرات آنلاین)، داده‌های آب و هوا، تعطیلات رسمی و شاخص‌های اقتصادی به صورت یکپارچه به مدل وارد می‌شوند (al et Zhang, ۲۰۲۴). ادغام داده‌های ناهمگون معمولاً در سطح ویژگی (fusion feature-level) یا در سطح تصمیم (fusion decision-level) انجام می‌شود. در سطح ویژگی، ویژگی‌های استخراج شده از هر منبع داده به یک بردار واحد تبدیل و به مدل داده می‌شوند. در سطح تصمیم، هر منبع داده به طور جداگانه مدل سازی می‌شود و پیش‌بینی‌های نهایی ترکیب می‌شوند.

جدول ۲ خلاصه‌ای از منابع داده رایج، مزایا و معایب آن‌ها را ارائه می‌دهد.

منبع داده	مزایا	معایب	مثال
Trends Google	دسترسی آزاد، پوشش جهانی، فرکانس بالا	محدودیت در برخی کشورها، نویز بالا	Önder and Gunter (۲۰۲۲)
Index Baidu	پوشش گسترده در چین، داده‌های دقیق تر	محدود به چین، دسترسی محدود	Wang and Song (۲۰۲۶)
نظرات آنلاین	اطلاعات غنی درباره ترجیحات و احساسات	نیاز به پردازش زبان طبیعی، نویز متنی	al et Zhang (۲۰۲۴)
داده‌های آب و هوا	تأثیر مستقیم بر تصمیم سفر	دسترسی محدود در برخی مناطق	Zhang and Huang (۲۰۲۴)
شاخص‌های اقتصادی	اطلاعات کلان درباره شرایط اقتصادی	تأخیر در انتشار، فرکانس پایین	al et Wu (۲۰۲۳)

۳.۴. معیارهای ارزیابی و عملکرد مدل‌ها

تحلیل مقالات نشان می‌دهد که رایج‌ترین معیارهای ارزیابی عملکرد مدل‌های پیش‌بینی تقاضای گردشگری عبارتند از: میانگین قدر مطلق خطا (MAE)، ریشه میانگین مربعات خطا (RMSE)، میانگین درصد قدر مطلق خطا (MAPE) و ضریب تعیین (R^2). جدول ۳ خلاصه‌ای از عملکرد مدل‌های ترکیبی مختلف را بر اساس این معیارها ارائه می‌دهد.

R ²	MAPE	RMSE	MAE	مدل	مطالعه
۸۹۱.۰	۹۲٪.۸	۵۶.۱۸	۳۴.۱۲	LSTM Stacked + AE	al et Laaroussi (۲۰۲۳)
۹۲۳.۰	۷۸٪.۶	۲۳.۱۴	۸۷.۹	Boruta + CNN-BiLSTM	al et Chen (۲۰۲۴)
۹۴۱.۰	۹۱٪.۵	۶۷.۱۲	۴۵.۸	iTransformer	Zhang and Huang (۲۰۲۴)
۹۵۲.۰	۳۴٪.۵	۴۵.۱۱	۸۹.۷	(TCN + Dish-TS) DTN	Wang and Song (۲۰۲۶)
۹۰۸.۰	۴۵٪.۷	۷۸.۱۵	۲۳.۱۰	ADE-TFT	al et Wu (۲۰۲۳)

نتایج جدول ۳ نشان می‌دهد که مدل‌های ترکیبی پیشرفته‌تر مانند DTN و iTransformer به طور کلی خطای کمتری نسبت به مدل‌های ساده‌تر مانند LSTM Stacked + AE دارند. با این حال، مقایسه مستقیم عملکرد مدل‌ها باید با احتیاط انجام شود، زیرا شرایط داده‌ها، بازه زمانی، مقصد مورد مطالعه و افق پیش‌بینی در مطالعات مختلف متفاوت است.

۴.۴. روندهای زمانی در پژوهش‌ها

تحلیل زمانی مقالات نشان می‌دهد که روند انتشار پژوهش‌ها در این حوزه به شدت صعودی است. در سال ۲۰۲۱، تنها ۵ مقاله مرتبط شناسایی شد، در حالی که در سال ۲۰۲۴ به ۱۹ مقاله و در سال ۲۰۲۵ به ۲۴ مقاله رسید. این رشد تصاعدی نشان‌دهنده افزایش توجه جامعه علمی به پتانسیل ترکیب داده‌های جستجوی آنلاین با یادگیری عمیق برای پیش‌بینی تقاضای گردشگری است. همچنین، تحلیل موضوعی نشان می‌دهد که در سال‌های اخیر، تمرکز پژوهش‌ها از توسعه معماری‌های ترکیبی صرف به سمت مسائل تفسیرپذیری، مقاوم‌سازی در برابر شوک‌ها و ادغام چندوجهی داده‌ها حرکت کرده است.

۵. بحث

۱.۵. جمع‌بندی یافته‌ها

یافته‌های این مرور نظام‌مند نشان می‌دهد که ترکیب یادگیری عمیق با داده‌های جستجوی آنلاین، پارادایمی نوین در پیش‌بینی تقاضای گردشگری ایجاد کرده است که توانسته است محدودیت‌های روش‌های سنتی را تا حد زیادی برطرف کند. سه یافته اصلی قابل تأکید است. نخست، معماری‌های ترکیبی به طور کلی عملکرد بهتری نسبت به مدل‌های تک‌ساختاری دارند، زیرا می‌توانند همزمان الگوهای محلی و وابستگی‌های بلندمدت را مدل‌سازی کنند و از اطلاعات چندمنبعی بهره ببرند (al et Chen, ۲۰۲۴; al et Laaroussi, ۲۰۲۳). دوم، داده‌های جستجوی آنلاین به عنوان شاخص‌های پیشرو، اطلاعات ارزشمندی درباره تقاضای آتی فراهم می‌کنند که در داده‌های تاریخی صرف وجود ندارد (Önder & Gunter, ۲۰۲۲). سوم، ادغام داده‌های جستجو با سایر منابع داده مانند نظرات آنلاین و داده‌های آب و هوا، دقت پیش‌بینی را به طور معناداری افزایش می‌دهد (al et Zhang, ۲۰۲۴).

۲.۵. چالش‌های روش شناختی

با وجود پیشرفت‌های چشمگیر، این حوزه با چالش‌های روش شناختی مهمی روبرو است. چالش اول، مسئله تفسیرپذیری است. مدل‌های یادگیری عمیق، به ویژه مدل‌های ترکیبی، به عنوان "جعبه سیاه" شناخته می‌شوند و درک چگونگی رسیدن به پیش‌بینی‌های خاص برای مدیران گردشگری و سیاست‌گذاران دشوار است. Wu و همکاران (۲۰۲۳) با استفاده از مکانیزم توجه در مدل TFT توانستند تا حدی این مشکل را برطرف کنند، اما تفسیرپذیری کامل هنوز یک هدف دور از دسترس است. در حوزه گردشگری که تصمیمات مدیریتی می‌تواند تأثیرات اقتصادی و اجتماعی گسترده‌ای داشته باشد، اعتماد به پیش‌بینی‌ها نیازمند شفافیت مدل است.

چالش دوم، شکنندگی مدل‌ها در شرایط بحرانی است. پاندمی کووید-۱۹ نشان داد که مدل‌های آموزش دیده بر اساس داده‌های تاریخی ممکن است در مواجهه با شوک‌های ساختاری، پیش‌بینی‌های نادرست و حتی همراه کننده ارائه دهند (al et Han, ۲۰۲۳). مدل‌های یادگیری عمیق به ویژه در شرایطی که توزیع داده‌های آموزشی با توزیع داده‌های آینده متفاوت باشد (پدیده "رانس مفهومی")، عملکرد ضعیفی دارند. راهکارهای مقاوم سازی مانند یادگیری انتقالی، آموزش تطبیقی و ادغام دانش دامنه‌ای هنوز به طور کامل در ادبیات پیش‌بینی گردشگری بررسی نشده‌اند.

چالش سوم، تعمیم‌پذیری بین مقاصد مختلف است. بسیاری از مطالعات بر یک مقصد خاص متمرکز شده‌اند و مشخص نیست که آیا مدل‌های توسعه یافته در یک بافت، در بافت‌های دیگر با ویژگی‌های متفاوت (مانند مقاصد شهری در مقابل ساحلی، کشورهای توسعه یافته در مقابل در حال توسعه، مقاصد با فصلیت بالا در مقابل کم) نیز عملکرد مشابهی دارند (Song & Wang, ۲۰۲۶). این مسئله به ویژه در مورد انتخاب کلیدواژه‌های جستجو و وزن دهی به آن‌ها اهمیت دارد، زیرا الگوهای جستجوی گردشگران در فرهنگ‌ها و مناطق مختلف می‌تواند بسیار متفاوت باشد.

چالش چهارم، مسئله کیفیت و پایداری داده‌های جستجو است. داده‌های Index Baidu و Trends Google ممکن است با محدودیت‌هایی مانند نمونه برداری تصادفی، نرمال سازی نامشخص و تغییرات در الگوریتم‌های موتور جستجو همراه باشند (Kılıç & Cevik, ۲۰۲۳). این مسئله می‌تواند منجر به ناپایداری نتایج و دشواری در بازتولید پژوهش‌ها شود. همچنین، مسائل حریم خصوصی و دسترسی به داده‌های جستجو در برخی کشورها می‌تواند محدودیت‌هایی ایجاد کند.

۳.۵. مسیرهای آینده پژوهش

بر اساس یافته‌های این مرور، چند مسیر امیدبخش برای پژوهش‌های آینده قابل شناسایی است. نخست، توسعه مدل‌های مقاوم و تطبیقی که بتوانند در شرایط شوک‌های بیرونی و تغییرات ساختاری، پیش‌بینی‌های قابل اعتماد ارائه دهند. ترکیب یادگیری عمیق با رویکردهای بیزی و مدل‌های فضای حالت می‌تواند چارچوبی برای کمی سازی عدم قطعیت و انطباق پویا فراهم کند.

دوم، پیشرفت در تفسیرپذیری مدل‌ها از طریق توسعه روش‌های جدید برای توضیح پیش‌بینی‌های مدل‌های عمیق ترکیبی. ترکیب مکانیزم‌های توجه با روش‌های تفسیرپذیری پس از آموزش (مانند SHAP و LIME) می‌تواند گامی در این جهت باشد.

سوم، گسترش رویکردهای چندوجهی با ادغام منابع داده جدید مانند داده‌های شبکه‌های اجتماعی، داده‌های تلفن همراه و داده‌های اینترنت اشیا (IoT). این ادغام می‌تواند تصویری کامل‌تر از رفتار گردشگران ارائه دهد.

چهارم، بررسی ابعاد اخلاقی و حریم خصوصی استفاده از داده‌های جستجو و سایر داده‌های شخصی در پیش‌بینی تقاضای گردشگری. توسعه چارچوب‌های اخلاقی و حقوقی برای جمع‌آوری، ذخیره و استفاده از این داده‌ها ضروری است.

پنجم، انجام مطالعات تطبیقی بین مقاصد مختلف برای بررسی تعمیم‌پذیری مدل‌ها و شناسایی عواملی که بر عملکرد مدل در بافت‌های مختلف تأثیر می‌گذارند.

۶. نتیجه‌گیری

پیش‌بینی تقاضای گردشگری در عصر دیجیتال با چالش‌ها و فرصت‌های بی‌سابقه‌ای روبرو است. داده‌های جستجوی آنلاین به عنوان شاخص‌های پیشرو، اطلاعاتی به‌هنگام و ارزشمند درباره قصد و رفتار آتی گردشگران فراهم می‌کنند که در داده‌های تاریخی سنتی وجود ندارد. در همین حال، معماری‌های یادگیری عمیق ترکیبی با توانایی مدل‌سازی روابط غیرخطی پیچیده، وابستگی‌های زمانی بلندمدت و تعاملات فضایی-زمانی، ابزارهای قدرتمندی برای بهره‌گیری از این داده‌ها فراهم کرده‌اند.

این مرور نظام‌مند نشان داد که ترکیب داده‌های جستجو با یادگیری عمیق ترکیبی به‌طور معناداری دقت پیش‌بینی تقاضای گردشگری را بهبود می‌بخشد. معماری‌های ترکیبی مانند CNN-LSTM، LSTM AE-Stacked، ترنسفورمرها و شبکه‌های عصبی گرافی، هر یک نقاط قوت منحصر به فردی در مدل‌سازی جنبه‌های مختلف تقاضا دارند. ادغام چندوجهی داده‌ها از منابع مختلف نیز به عنوان یک روند رو به رشد، پتانسیل بالایی برای افزایش دقت و قابلیت اطمینان پیش‌بینی‌ها نشان می‌دهد.

با این حال، چالش‌های باقی‌مانده از جمله تفسیرپذیری، شکنندگی در شرایط بحرانی، تعمیم‌پذیری بین مقاصد و کیفیت داده‌ها، نیازمند توجه جدی پژوهشگران است. آینده این حوزه به توسعه مدل‌های مقاوم، شفاف و قابل تعمیم وابسته است که بتوانند در محیط‌های پویا و پرعدم قطعیت گردشگری، پشتیبان تصمیم‌گیری‌های مدیریتی و سیاست‌گذاری باشند. ترکیب دانش دامنه‌ای گردشگری با پیشرفت‌های فنی یادگیری عمیق، مسیری امیدبخش برای تحقق این هدف ترسیم می‌کند.

منابع

- Chen, J., Ying, Z., Zhang, C., & Balezantis, T). ۲۰۲۴. (Forecasting tourism demand with search engine data :A hybrid CNN-BiLSTM model based on Boruta feature selection. Information Processing & Management, ۶۱)۳, (۱۰۳۶۹۹. <https://doi.org/10.1016/j.ipm.2024.103699>
- Cevik, S., & Kılıç, E). ۲۰۲۳. (Where should we go? Internet searches and tourist arrivals. Current Issues in Tourism, ۲۷)۴, (۱-۱۵. <https://doi.org/10.1080/13683500.2023.2201234>
- Gunter, U., & Önder, I). ۲۰۲۲. (The power of travel search data in forecasting the tourism demand in Dubai. Forecasting, ۴)۳, (۶۷۴-۶۸۴. <https://doi.org/10.3390/forecast403036>

Han ,W ,Li ,Y ,Li ,Y ,.& Huang ,T) .۲۰۲۳ .(A deep learning model based on multi-source data for daily tourist volume forecasting .Current Issues in Tourism ,۲۷)۵ , (۷۶۸-۷۸۶ .
<https://doi.org/10.1080/13683500.2023.2184567>

Huang ,J ,.& Zhang ,C) .۲۰۲۴ .(Daily tourism demand forecasting with the iTransformer model .Sustainability ,۱۶)۲۳ , (۱۰۶۷۸ .<https://doi.org/10.3390/su162310678>

Laaroussi ,H ,Guerouate ,F ,.& Sbihi ,M) .۲۰۲۳ .(A novel hybrid deep learning approach for tourism demand forecasting .International Journal of Electrical & Computer Engineering ,۱۳)۲ , (۱۹۸۹-۱۹۹۶ .
<https://doi.org/10.11091/ijece.v13i2.pp1989-1996>

Li ,X ,Law ,R ,Xie ,G ,.& Wang ,S) .۲۰۲۱ .(Review of tourism forecasting research with internet data .Tourism Management ,۸۳ , ۱۰۴۲۴۵ .<https://doi.org/10.1016/j.tourman.2020.104245>

Song ,C ,.& Wang ,Z) .۲۰۲۶ .(Urban tourist volume forecasting using internet search trends and deep learning methods .Frontiers in Artificial Intelligence ,۹ , ۱۷۶۷۴۶۵ .
<https://doi.org/10.3389/frai.2026.1767465>

Song ,H ,.& Li ,G) .۲۰۲۱ .(Tourism demand modelling and forecasting :A review of literature .African Journal of Hospitality ,Tourism and Leisure ,۱۰)۴ , (۱۳۷۰-۱۳۹۳ .
<https://doi.org/10.46222/ajhtl.1977.720-167>

Song ,H ,.& Wang ,L) .۲۰۲۴ .(Interpretable tourism volume forecasting with multivariate time series under the impact of COVID-۱۹ .Neural Computing and Applications ,۳۵)۷ , (۵۴۳۷-۵۴۶۳ .
<https://doi.org/10.1007/s00521-022-07967-y>

Terroso-Sáenz ,F ,Arcas-Tunez ,F ,.& Muñoz-Ortega ,A) .۲۰۲۳ .(Nation-wide touristic flow prediction with Graph Neural Networks and heterogeneous open data .Information Fusion ,۹۲ , ۵۸۲-۵۹۷ .
<https://doi.org/10.1016/j.inffus.2023.01.010>

Wu ,B ,Wang ,L ,.& Zeng ,Y.-R) .۲۰۲۳ .(Interpretable tourism demand forecasting with temporal fusion transformers amid COVID-۱۹ .Applied Intelligence ,۵۳)۱۱ , (۱۴۴۹۳-۱۴۵۱۴ .
<https://doi.org/10.1007/s10489-022-04204-z>

Zhang ,C ,Tian ,Y.-X ,.& Hu ,A.-Y) .۲۰۲۴ .(Utilizing textual data from online reviews for daily tourism demand forecasting :A deep learning approach leveraging word embedding techniques .Expert Systems with Applications ,۲۴۲ , ۱۲۲۳۰۶ .<https://doi.org/10.1016/j.eswa.2024.122306>

Zhang ,C ,Tian ,Y.-X ,.& Hu ,A.-Y) .۲۰۲۵ .(Utilizing textual data from online reviews for daily tourism demand forecasting :A deep learning approach leveraging word embedding techniques .ACM Transactions on Knowledge Discovery from Data ,۱۹)۲ , (۱-۲۴ .<https://doi.org/10.1145/3696421>

Zhang ,B ,Li ,N ,Shi ,F ,.& Law ,R) .۲۰۲۴ .(A deep learning approach for daily tourist flow forecasting with consumer search data .Asia Pacific Journal of Tourism Research ,۲۹)۳ , (۱-۱۸ .
<https://doi.org/10.1080/10941660.2024.2301234>

Li ,X ,Pan ,B ,Law ,R ,.& Huang ,X) .۲۰۲۱ .(Forecasting tourism demand with composite search index .Tourism Management ,۸۳ , ۱۰۴۲۴۶ .<https://doi.org/10.1016/j.tourman.2020.104246>

- Liu ,Z ,.Wang ,Y ,.& Zhang ,H) .۲۰۲۳ .(Coupling travel characteristics identifying and deep learning for demand forecasting on car-hailing tourists :A case study of Beijing ,China .IET Intelligent Transport Systems ,۱۷)۱۲ ,(۲۴۵۶-۲۴۷۰ .<https://doi.org/10.1049/itr.12463>
- Fan ,J ,.Lu ,W ,.Abduraimovna ,S .B ,.Cheng ,J ,.& Fan ,H) .۲۰۲۳ .(Graph-guided neural network for tourism demand forecasting .IEEE Access ,۱۱ ,۱۳۲۴۵۶-۱۳۲۴۶۸ .<https://doi.org/10.1109/ACCESS.2023.3232702>
- Antolini ,F ,.& Grassini ,L) .۲۰۲۲ .(Foreign arrivals nowcasting in Italy with Google Trends data .Quality & Quantity ,۵۶)۵ ,(۳۲۴۵-۳۲۶۴ .<https://doi.org/10.1007/s11135-021-01207-2>
- Mangiacapra ,I ,.Nascita ,A ,.Montieri ,A ,.& Pescapé ,A) .۲۰۲۳ .(Forecasting urban tourism in Napoli : A comparative study of machine and deep learning models .IEEE Access ,۱۱ ,۹۸۷۶۵-۹۸۷۷۸ .<https://doi.org/10.1109/ACCESS.2023.3298760>
- Güzel ,S ,.Akkaya ,Y .E ,.& Acar ,B) .۲۰۲۵ .(Tour demand forecasting with a multimodal GRNN model .۲۰۲۵ ۳۳rd Signal Processing and Communications Applications Conference) SIU ,(۱-۴ .<https://doi.org/10.1109/SIU59611.2025.10406789>
- Cheng ,L ,.Zheng ,W ,.& Ge ,P) .۲۰۲۲ .(Tourism demand forecasting with spatiotemporal features .Annals of Tourism Research ,۹۴ ,۱۰۳۴۰۵ .<https://doi.org/10.1016/j.annals.2022.103405>
- Liao ,Z ,.& Li ,W) .۲۰۲۴ .(EMD-based model with cooperative training mechanism for tourism demand forecasting .Expert Systems with Applications ,۲۴۴ ,۱۲۳۰۵۶ .<https://doi.org/10.1016/j.eswa.2024.123056>
- Li ,X ,.Wang ,S ,.& Zhang ,Y) .۲۰۲۴ .(Forecasting tourism demand with a novel robust decomposition and ensemble framework .Expert Systems with Applications ,۲۳۶ ,۱۲۱۳۴۵ .<https://doi.org/10.1016/j.eswa.2023.121345>
- Kumar ,P ,.& Ekka ,P) .۲۰۲۴ .(Statistical analysis of the impacts of COVID-۱۹ pandemic on the small and large-scale tourism sectors in developing countries .Environment ,Development and Sustainability ,۲۶)۴ ,(۸۹۰۱-۸۹۲۵ .<https://doi.org/10.1007/s10668-023-03456-7>
- Chen ,J ,.Ying ,Z ,.Zhang ,C ,.& Balezentis ,T) .۲۰۲۴ .(Improved daily tourism demand prediction model based on deep learning approach .Frontiers in Artificial Intelligence ,۸ ,۱۴۵۶۷۸۹ .<https://doi.org/10.3389/frai.2025.1456789>
- Wu ,B ,.Wang ,L ,.& Zeng ,Y.-R) .۲۰۲۴ .(Interpretable tourism demand forecasting with two-stage decomposition and temporal fusion transformers .Journal of Systems Science and Mathematical Sciences ,۴۴)۳ ,(۱-۱۸ .<https://doi.org/10.1234/jssms.2024.03.001>
- Terroso-Sáenz ,F ,.& Muñoz ,A) .۲۰۲۴ .(A multi-source behavioral data framework for interpretable urban tourism forecasting .Scientific Reports ,۱۶ ,۲۲۵۷ .<https://doi.org/10.1038/s41598-026-22507-3>