

توسعه یک سیستم توصیه‌گر برای هدایت گردشگران در سفرهای کوتاه شهری

ارسلان رضوی^۱ | فرهاد حسینی^{۲*}

۱ دانش‌آموخته کارشناسی ارشد مهندسی نقشه‌برداری-سیستم‌های اطلاعات مکانی، دانشکده مهندسی عمران، دانشگاه تربیت دبیر شهید رجایی، تهران

۲ استادیار گروه مهندسی نقشه برداری، دانشکده مهندسی عمران، دانشگاه تربیت دبیر شهید رجایی، تهران

اطلاعات مقاله

چکیده

سابقه مقاله

دریافت مقاله: ۱۴۰۰/۰۸/۲۹

پذیرش مقاله: ۱۴۰۰/۱۲/۰۱

کلمات کلیدی

سیستم توصیه گر، گردشگری
شهری، نقاط مورد علاقه،
یادگیری ماشین

انتخاب مکان مناسب برای تفریح، گردش و ورزش سؤال هر روزه و مهمی برای بسیاری از مردم است. سؤالاتی مانند اینکه چه مراکز تفریحی و ورزشی متناسب با بودجه و زمان گروه‌های مختلف مردم است، همواره وجود دارد. سیستم‌های توصیه‌گر متنوعی برای پیشنهاد بهترین گزینه‌ها به کاربران وجود دارد. در این تحقیق با استفاده الگوریتم‌های مختلف یادگیری ماشین سعی شده است تا بهترین پیشنهادها به کاربران توسط یک سیستم توصیه‌گر چند منظوره و با توجه به سلايق و سوابق آن‌ها ارائه گردد. در این پژوهش بر خلاف تحقیقات پیشین که سیستم‌های توصیه‌گر با توجه به برنامه‌های بلندمدت یا میان‌مدت چند روزه و هفتگی اقدام به ارائه پیشنهاد محل تفریحی می‌کردند؛ هدف، ارائه پیشنهادهایی برای کاربران در بازه‌های کوتاه چند ساعته است. برای اینکه سیستم توصیه‌گر بتواند تشخیص دهد که به کاربر پیشنهاد محل تفریحی و ورزشی روباز یا سرپوشیده را بدهد، از الگوریتم ماشین‌های بردار پشتیبان و استفاده شده است. همچنین برای ارائه پیشنهادهایی متناسب با بودجه و زمان کاربر، روش خوشه‌بندی K میانگین به خدمت گرفته شده است. در انتها پیش‌بینی میزان رضایت کلی کاربر نسبت به رضایت از محل پیشنهادی و نیز نسبت به مسیر پیشنهادی بررسی می‌شود. این سیستم توسط کاربران در محدوده مرکزی شهر تهران به خدمت گرفته شد و نتایج کمی نشان داد که به صورت کلی رضایت کاربران از مقصد پیشنهادی اهمیت بیشتری در رضایت کلی کاربران دارد (حدود ۱/۲ برابر) تا رضایت کاربران نسبت به مسیر پیشنهادی.



۱ مقدمه

مردم روزانه زمان و پول زیادی را صرف تفریح و گردش در نقاط مختلف شهری می‌کنند. افرادی که قصد گردش و تفریح دارند، یا به دنبال مراکز تفریحی و فرهنگی در شهر خود هستند یا در شهری دیگر به عنوان گردشگر به جست‌وجوی آن می‌پردازند. حجم بسیار زیادی از اطلاعات در اینترنت موجود است، بنابراین فرآیند تصمیم‌گیری با استفاده از این اطلاعات، آهسته، پیچیده و زمانبر است. سیستم‌های توصیه‌گر مشکل انباشتگی فراوان اطلاعات را کم می‌کنند و با استفاده از ترجیحات کاربران، علایق و موقعیت آن‌ها، پیشنهادهایی را به کاربران ارائه می‌دهند [۱]. سیستم‌های توصیه‌گر یکی از انواع سیستم‌های یادگیری ماشین هستند که با تحلیل رفتار کاربر خود، اقدام به پیشنهاد مناسب‌ترین اقلام (داده، اطلاعات، کالا و غیره) به وی می‌نماید. این نوع سیستم رویکردی است که برای مواجهه با مشکلات ناشی از حجم فراوان و رو به رشد اطلاعات، ارائه شده است و به کاربر خود کمک می‌کند تا در میان حجم عظیم اطلاعات، سریع‌تر به هدف خود نزدیک شود. اولین تجربه موفق سیستم‌های توصیه‌گر در وبسایت‌های تجاری بود که اطلاعاتی را درباره محصولات که احتمال می‌رفت مورد توجه کاربران قرار گیرد، ارائه می‌دادند که شامل اقلامی همچون فیلم، کتاب، اخبار و صفحات وب بودند. اخیراً سیستم‌های توصیه‌گر به وفور در زمینه گردشگری الکترونیکی مورد استفاده قرار گرفته‌اند. این سیستم‌ها به ارائه خدماتی مانند توصیه در گردش و فعالیت‌ها و یافتن نقاط مورد علاقه و پکیج‌های گردشگری که متناسب با ترجیحات کاربر باشد، می‌پردازند [۲-۵]. گردشگران معمولاً برای محل اقامت، غذا، حمل‌ونقل یا هزینه ورودی مکان‌های دیدنی، محدودیت‌های مالی دارند. بنابراین محدودیت‌های مالی هم در کنار محدودیت‌های زمانی، گزینه‌های مناسب برای کاربر را کاهش می‌دهد. تا پیش از این سیستم‌های توصیه‌گر یا به معرفی محل‌های دیدنی می‌پرداختند یا هتل‌ها و رستوران‌ها را پیشنهاد می‌دادند، اما امروزه نیازمند سیستمی هستیم که بتواند این دو قسمت را ادغام کند. دلیل این امر این است که این دو عامل بر یکدیگر تأثیر مستقیم و غیر مستقیم دارند. به طور مثال محل هتل بر روی انتخاب جاهای دیدنی تأثیر دارد [۶]. بسیاری از سیستم‌های توصیه‌گر حول استفاده از یادگیری ماشین و الگوریتم‌های داده کاوی قرار گرفته‌اند تا بتوانند امتیاز و نظر کاربر به محصول یا ترتیب پیشنهاد محصولات به کاربر را پیش بینی کنند [۷]. وجود سیستمی که بتواند عوامل مختلف و ویژگی‌های متنوع را در نظر بگیرد و بهترین پیشنهاد را به صورت آنی به کاربر ارائه دهد می‌تواند موجب صرفه جویی در وقت، هزینه و همچنین تشویق مردم به تفریح بیشتر شود که موجب افزایش نشاط در سطح جامعه می‌شود. همچنین با از بین بردن یا

کاهش نیاز به تورهای گردشگری و نیاز به سؤال پرسیدن از مردم محلی باعث رونق توریسم خارجی در آن شهر و به تبع آن موجب رونق اقتصاد محلی می‌شود. راه‌اندازی سیستم مذکور نیازمند حجم بالایی از اطلاعات است که برای استفاده از این اطلاعات نیازمند آشنایی با علوم نوظهور داده و یادگیری ماشین هستیم. به کمک الگوریتم‌های مختلف یادگیری ماشین می‌توان بهترین پیشنهادها را به عنوان اولین پیشنهاد به کاربر ارائه داد، رفتار آینده مردم را پیش‌بینی کرد و سنجید و تدابیر لازم برای برطرف کردن نیازهای آن‌ها اندیشید [۸]. همچنین با تحلیل نتایج حاصل از تحلیل‌های علوم داده و یادگیری ماشین می‌توان به تدریج عملکرد سیستم یا مدل را بهبود داد تا به بهترین نتایج مورد پسند کاربران رسید. مسئله اصلی در این تحقیق یافتن مناسب‌ترین مقصد و مسیر برای مقاصد گوناگون از جمله تفریحی و فرهنگی و ورزشی و غیره با توجه به موقعیت مکانی فرد، محدودیت‌های زمانی کاربر، میزان تمایل فرد در سپری کردن مدت زمانی در راه، پیاده یا سواره بودن کاربر و ترجیح فرد به استفاده از نوع خاصی از خدمات است [۹].

۲- پیشینه تحقیق

سیستم‌های توصیه‌گر در حوزه گردشگری به کاربران کمک می‌کنند تا به مناسب‌ترین گزینه دست یابند. بعضی از سیستم‌های توصیه‌گر گردشگری متمرکز بر جاذبه‌ها و مقاصد گردشگری هستند و دسته‌ای دیگر از آن‌ها به پیشنهاد برنامه تور که شامل حمل و نقل، رستوران و وسایل راحتی می‌شود، می‌پردازند [۱۰]. چندین سیستم پیشنهاددهنده رستوران تاکنون عرضه شده است. برای مثال بورک و همراهان سیستم توصیه‌گری با نام Entrée پیشنهاد دادند که به پیشنهاد رستوران با استفاده از روش KB می‌پرداخت. دانش مورد نیاز با استفاده از کاربران جمع‌آوری می‌شد و توسط Entrée بازیابی می‌شد تا با پالایش عواملی همچون قیمت و مزه، گزینه‌های مشابه را پیدا کنند [۱۱]. بورک با ترکیب کردن CF با KB توانست Entrée را ارتقاء دهد، به این معنی که جدا از ویژگی‌های رستوران‌ها، ارزیابی کاربران نیز مد نظر قرار گرفت [۱۲]. گوشی‌های همراه امکان توسعه سیستم‌های توصیه‌گر بر مبنای گوشی‌های همراه را فراهم می‌آورند. ون و وون سیستمی را برای پیشنهاد رستوران در تایپه پایتخت تایوان برای گردشگران طراحی کردند. این سیستم یک سیستم توصیه‌گر بر مبنای مشارکت است که به کاربران این اجازه را می‌دهد پیشنهادهایی آنی از یک برنامه کاربردی تلفن همراه دریافت کنند [۱۳]. کتیس یک سیستم توصیه‌گر بافت‌آگاه است که رستوران، جاذبه‌های گردشگری و محل سکونت را به گردشگران پیشنهاد می‌دهد. اطلاعات زمینه‌ای مانند محل از طریق مدیر زمینه‌ای جمع‌آوری می‌شود. از مجموعه‌ای از خدمات تحت

وب که به‌وسیله یک برنامه کاربردی کارگذار سرور تهیه می‌شود، استفاده می‌گردد تا اطلاعات زمینه‌ای کاربر را جمع‌آوری کند. با ترکیب پرسش از کاربر و اطلاعات زمینه‌ای کاربر که از برنامه کارگذار تهیه می‌شود، توصیه‌ها و پیشنهادها به کاربر داده می‌شوند [۱۴]. رجا یک سیستم توصیه‌گر برای پیدا کردن رستوران است. این سیستم روش‌های CF و KB را ترکیب می‌کند. در این سیستم چنانچه بتوان پروفایل کاربری بر اساس امتیاز دهی‌های کاربر تشکیل داد، از روش CF برای پیشنهاد دادن استفاده می‌شود. در شرایطی که سیستم اطلاعات ناکافی درباره کاربر دارد، از یک روش استدلال موردی استفاده می‌شود [۱۵]. پیسیس به گردشگران در شهر اُپرتو پرتغال کمک می‌کند تا برنامه تفریحی شخصی‌سازی شده خود را بیابند. این سیستم از یک روش توصیه ترکیبی استفاده می‌کند. این سیستم از CF و CB به همراه یک روش خوشه‌بندی و یک الگوریتم کلاس‌بندی شرکت‌پذیر، استفاده می‌کند و همچنین منطق فازی را برای بهبود کیفیت توصیه‌ها به کار می‌گیرد [۱۶]. SigTur/E-Destination برای تولید پیشنهادهای شخصی برای گردشگران در منطقه تاراگنا در اسپانیا طراحی شده است. برای رسیدن به پیشنهادهای مناسب، SigTur/E-Destination چندین نوع از اطلاعات و روش‌های توصیه را ترکیب می‌کند. اطلاعاتی که در این سیستم استفاده می‌شود شامل داده‌های جمعیتی، جزئیاتی که محتوای سفر را مشخص می‌کند، جنبه‌های جغرافیایی، اطلاعات تهیه شده به صورت صریح و واضح به وسیله کاربر و بازخوردهای ضمنی ناشی از تعامل کاربر و سیستم است. SigTur/E-Destination از چندین تکنیک توصیه استفاده می‌کند، از جمله از قالب‌ها، و تکنیک‌های CB و CF و ابزارهای هوش مصنوعی شامل الگوریتم‌های خودکار خوشه‌بندی، مدیریت هستی‌شناسی و تعریف اندازه‌گیری جدید شباهت بین کاربران، بر اساس اپراتورها پیچیده تجمیع [۱۷]. SMARTMUSEUM یک سیستم توصیه‌گر تلفن همراه است که محل‌هایی را بر روی گوشی به کاربر پیشنهاد می‌دهد. در این سیستم یک شخصی‌سازی هستی‌شناسانه، تفسیر و چارچوب پالایش اطلاعات، توسعه داده شده است. اطلاعات زمینه‌ای یا به صورت ورودی توسط کاربر وارد می‌شود یا با استفاده از سنجنده‌های تعبیه شده در داخل تلفن همراه به دست می‌آیند و بر اساس مفاهیم تعیین شده در هستی‌شناسی، تصویر می‌شوند [۱۸]. iTravel یک سیستم دیگر توصیه‌گر گردشگری بر پایه تلفن همراه است. در این سیستم تکنیک‌های CF و ارتباط هم‌تا به هم‌تا بر پایه تلفن همراه ترکیب شده‌اند. برای استفاده از اطلاعات سایر کاربران با علایق مشابه در استفاده از موبایل در گردشگری، سه راه تبادل داده برای کاربران برای تبادل امتیازاتشان به جاذبه‌هایی که دیدن نموده‌اند، پیشنهاد داده شده است [۱۹]. Moleskiing یک وب‌سایت برای کمک به مردم جهت برنامه‌ریزی برای اسکی کردن

است. این سیستم توصیه‌گر به کاربران اجازه می‌دهد تا تجربیات و نظراتشان را نسبت به محلی خاص به اشتراک بگذارند. در این سیستم هر کاربر درجه‌ای از اطمینان دارد. مردمی که قصد اسکی رفتن دارند، می‌توانند از اطلاعات وضعیت برف استفاده کنند تا به یک مسیر شخصی‌سازی شده و مطمئن برسند [۲۰]. DIETORECS یک سیستم توصیه‌گر استدلال موردی است که برنامه‌ای کامل را برای گردشگران فراهم می‌کند. کاربران می‌توانند با توجه به تجربیاتشان، از سیستم به روش‌های مختلفی استفاده کنند. کاربران با تجربه می‌توانند ترجیحات ریز و دقیقی را برای نقاط مورد علاقه خود درست کنند، در حالی که کاربران با تجربیات کمتر تنها می‌توانند لیستی از نقاط مورد علاقه خود درست کنند [۲۱]. MAS-TROCARONTE یک سیستم توصیه‌گر همراه برای رانندگان است که از روش‌های KB برای پیشنهاد جاذبه‌ها، رستوران‌ها و هتل‌ها استفاده می‌کند. این سیستم توصیه‌گر از اطلاعات زمینه‌ای برای پیشنهاد گزینه‌های مناسب برای رانندگان استفاده می‌کند، برای مثال رستوران‌هایی که در آن ساعت خاص قادر به ارائه سرویس هستند و یا ایستگاه‌های سوخت نزدیک را هنگامی که خودرو نیاز به سوخت‌گیری دارد، پیشنهاد می‌دهد [۲۲]. SPETA یک سیستم توصیه‌گر است که از اطلاعات موقعیت فعلی کاربر، ترجیحات و سابقه محل‌های قبلی کاربر استفاده می‌کند تا سرویس‌هایی را پیشنهاد دهد که گردشگران از راهنماهای سفر انتظار دارند. این سیستم، شبکه‌های اجتماعی، وب معنایی و متن-آگاهی در سیستم‌های فراگیر را ترکیب می‌کند تا بتواند تجربیات گردشگران را بهبود بخشد. این سیستم راهنمایی شخصی شده‌ای را ارائه می‌دهد و با استفاده از معنا، موقعیت جغرافیایی و تکنیک‌های اجتماعی، مشکل عدم یکپارچگی سرویس به گردشگران در رابطه با جست‌وجو، پیدا کردن و ارائه سرویس‌های شخصی را حل می‌کند [۲۳]. Traveller سیستمی است که به پیشنهادهایی در رابطه با تعطیلات و تورها می‌پردازد. این سیستم عاملی را می‌سازد که CF را با CB و روش‌های توصیه وابسته به جمعیت ترکیب می‌کند [۲۴]. با توجه به میزان پیچیدگی و پیشنیازهای آیت‌های پیشنهادی، سیستم‌های توصیه‌گر متنوعی در گردشگری مورد استفاده قرار می‌گیرد. برای پیشنهاد آیت‌های به نسبت ساده‌ای مانند رستوران‌ها، معمولاً تکنیک‌های CF و CB مورد استفاده قرار می‌گیرند. برای پیشنهاد آیت‌هایی که بیشتر پیچیده هستند مانند مسیرهای گردش و برنامه‌ریزی زمانی، KB و تکنیک‌های ترکیبی توصیه‌ای با حوزه‌ای از دانش مورد استفاده قرار می‌گیرند. برای پیشنهاد کردن آیت‌هایی با نیازهای آنی، مانند ایستگاه‌های سوخت، روش‌های بافت‌آگاه مورد استفاده قرار می‌گیرند [۱۰]. بیان و هوانگ در سال ۲۰۰۹ میلادی سامانه هوشمندی را برای دریافت پیشنهادهای شخصی از جاذبه‌های گردشگری در یک شهر طراحی کرده‌اند

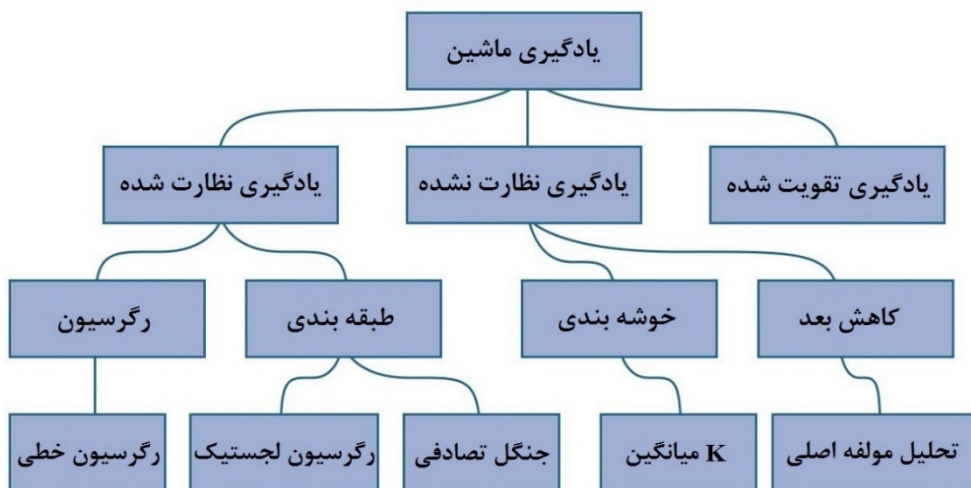
که معماری آن سه لایه است. نخستین لایه، مرورگر وب است که کاربران از طریق آن درخواست‌هایشان را ارسال و نتایج پیشنهادی را دریافت می‌کنند. لایه میانی، دو نوع سرور (سرور نقشه وب و وب سرور) را شامل می‌شود که در آن اطلاعات برخط جاذبه‌های گردشگری که فراهم‌کنندگان اطلاعات مسافرتی گوناگون آن‌ها را گردآوری کرده‌اند، در هستی‌شناسی توریسم بر مبنای Web Ontology Language نمایش داده می‌شود. لایه سوم شامل وب سرورهای مکانی از قبیل Arc web services است [۲۵]. سامانه دیگری در این زمینه I'm felling Loco است که در سال ۲۰۱۲ میلادی توسط Savage و همکاران ارائه شد که سامانه توصیه‌گر مکان-مبنای همه جا آگاهی است که به صورت خودکار اولویت‌های کاربر و محدودیت‌های زمان و مکان را برای پیشنهاد مکان‌ها در بر می‌گیرد. سامانه از طریق داده‌کاوی پروفایل مشخص در چارچوب شبکه اجتماعی مکان-مبنا، اولویت‌های کاربر را استخراج می‌کند [۲۶]. سیستم Cyberguide نقشه راهنما به همراه اطلاعات کلی سفر و سیستم Travel Planner اطلاعات نوع پرواز، محل اقامت، رستوران و نقاط دیدنی را به کاربر ارائه می‌دهد. همچنین وب‌سایت گردشگری TripAdvisor به هر کاربر سفرها، مکان‌ها و فعالیت‌هایی را پیشنهاد می‌کند و افزون بر آن شامل یک مؤلفه اجتماعی است که به کاربران امکان بررسی، اظهار نظر و امتیازدهی به تعداد زیادی از عناصر را می‌دهد و به فرآیند پیچیده تصمیم‌گیری در حیطه گردشگری کمک می‌کند. تفاوت اصلی مدل پیشنهاد شده در این پژوهش با مدل‌های ارائه شده قبلی، برنامه ریزی کوتاه‌مدت ساعتی برای یک روز است. مدل‌های قبلی بر مبنای برنامه‌ریزی چند روز در یک هفته یا روزهای یک ماه بودند. همچنین در این مدل به عامل مهم هزینه نیز توجه شده است که در مدل‌های قبلی اغلب غایب بوده است. بخشی از داده‌های مورد نیاز پژوهش نیز به صورت داوطلبانه توسط کاربران ارائه می‌شود. همچنین عواملی همچون شرایط آب و هوایی و ساعات کاری نیز در نظر گرفته می‌شوند.

۳- مواد و روش‌ها

یادگیری ماشین به معنی مطالعه علمی الگوریتم‌ها و مدل‌های آماری مورد استفاده در سیستم‌های کامپیوتری برای انجام وظیفه و کاری مشخص با تکیه بر الگوها و استنتاج، بدون دریافت مستقیم دستورالعمل از انسان است. یادگیری ماشین یکی از زیر شاخه‌های هوش مصنوعی است. الگوریتم‌های یادگیری ماشین، با استفاده از داده‌های نمونه که به عنوان داده‌های آموزشی نیز شناخته می‌شوند، مدل‌های ریاضی را می‌سازند تا به پیش‌بینی و تصمیم‌بپردازند بدون اینکه به صورت صریح برای انجام آن برنامه‌ریزی شده باشند [۲۷، ۲۸]. الگوریتم‌های یادگیری ماشین در موارد بسیاری کاربرد دارند، از جمله

فیلتر کردن ایمیل و بینایی ماشین که با استفاده از دستورالعمل‌های مشخص، امکانپذیر نیست. یادگیری ماشین به آمار محاسباتی بسیار نزدیک است. داده‌کاو نیز یک حوزه مطالعاتی در یادگیری ماشین است که به کاوش و آنالیز داده در یادگیری نظارت نشده می‌پردازد [۲۸، ۲۹].

یادگیری نظارت شده یک روش عمومی در یادگیری ماشین است که در آن به سیستم، مجموعه‌ای از جفت‌های ورودی خروجی ارائه شده و سیستم تلاش می‌کند تا تابعی از ورودی به خروجی را فرا گیرد. یادگیری تحت نظارت نیازمند تعدادی داده ورودی به منظور آموزش سیستم است [۳۰]. مثال‌های زیادی در یادگیری ماشینی وجود دارد که در دسته یادگیری با نظارت قرار می‌گیرد، از جمله درخت تصمیم‌گیری، ماشین‌های بردار پشتیبان، رگرسیون خطی، رگرسیون لجستیک، شبکه‌های عصبی مصنوعی و بسیاری نمونه‌های دیگر. یادگیری نظارت نشده در مقابل یادگیری نظارت شده قرار می‌گیرد و یکی از انواع یادگیری در یادگیری ماشینی است. اگر یادگیری بر روی داده‌های بدون برچسب و برای یافتن الگوهای پنهان در این داده‌ها انجام شود، یادگیری بدون نظارت [۳۱] خواهد بود. از انواع یادگیری بدون نظارت می‌توان به خوشه‌بندی، کاهش ابعاد و روش‌های تشخیص ناهنجاری اشاره کرد [۳۲، ۳۳]. نمودار سلسله مراتبی انواع الگوریتم‌های یادگیری ماشین در شکل ۱ نمایش داده شده است.



شکل ۱: نمودار سلسله مراتبی انواع الگوریتم‌های یادگیری ماشین [۳۳]

در یادگیری ماشین و آمار، طبقه‌بندی مسئله شناسایی تعلق یک مشاهده جدید به یک دسته از دسته‌ها (زیر جمعیت‌ها)، بر اساس یک مجموعه از داده‌های مورد استفاده به منظور آموزش شامل مشاهدات است که عضویت در دسته‌هایشان معلوم است. الگوریتم‌هایی چون درخت تصمیم، K نزدیک‌ترین همسایه، رگرسیون لجستیک و شبکه عصبی، نمونه‌هایی از الگوریتم‌های طبقه‌بندی هستند [۳۴].

۱-۳- رگرسیون

در مدل‌های آماری، تحلیل رگرسیون یک فرآیند آماری برای تخمین روابط بین متغیرها است. این روش شامل تکنیک‌های زیادی برای مدل‌سازی و تحلیل متغیرهای خاص و منحصر به فرد، با تمرکز بر رابطه بین متغیر وابسته و یک یا چند متغیر مستقل است. تحلیل رگرسیون خصوصاً کمک می‌کند در فهم اینکه چگونه مقدار متغیر وابسته با تغییر هر یک از متغیرهای مستقل و با ثابت بودن دیگر متغیرهای مستقل تغییر می‌کند. چنانچه بیش از یک متغیر مستقل در محاسبات وجود داشته باشد، رگرسیون را چند متغیره می‌نامند. تحلیل رگرسیون به صورت گسترده برای پیش‌بینی و نیز برای شناخت ارتباط میان متغیر مستقل و وابسته و شکل این روابط استفاده می‌شود. تحلیل رگرسیون به دو دسته خطی و غیر خطی تقسیم می‌شود که خطی یا غیر خطی بودن رگرسیون، وابسته به نوع رابطه بین متغیرهای وابسته و مستقل است. در رگرسیون خطی، متغیر وابسته y ترکیبی خطی از ورودی یا متغیرهای مستقل است ولی در رگرسیون غیر خطی چنین نیست [۳۵]. بدون انجام یک نمونه‌گیری بسیار بزرگ، تحلیل رگرسیون تنها می‌تواند به صورت تقریبی، تأثیر متغیر ورودی بر خروجی را برآورد نماید و میزان این تأثیر برآورد شده می‌تواند کوچکتر یا بزرگتر از مقدار واقعی آن باشد. بنابراین خوب است که بدانیم دامنه این تغییرات احتمالاً چقدر است. به بیان دیگر احتمال این وجود دارد که ضریب رگرسیون حداقل و حداکثر چقدر باشد. برای رسیدن به این هدف، از مفهومی به نام بازه اطمینان استفاده می‌کنیم. بازه اطمینان برای شیب خط رگرسیون یا ضریب رگرسیون با فرض معادله خط رگرسیون به صورت $Y = \beta_0 + \beta_1 X$ با استفاده از روابط زیر محاسبه می‌شود:

$$\hat{\beta}_1 - t_{\alpha/2, df} \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \leq \beta_1 \leq \hat{\beta}_1 + t_{\alpha/2, df} \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \quad (1)$$

$$S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2)$$

$$\hat{\sigma}^2 = MSE = \frac{\sum_{i=1}^n r_i^2}{n-2} \quad (3)$$

$$r_i = y_i - \hat{y}_i \quad (4)$$

$$n - (k + 1) \quad (5)$$

مقدار t با توجه بازه اطمینان انتخاب شده مثلاً ۹۵ درصد و درجه آزادی با استفاده از جدول توزیع t تعیین می‌شود. درجه آزادی نیز با استفاده از رابطه ۵ محاسبه می‌شود که n تعداد نمونه‌ها و k تعداد متغیرهای مستقل است که در رگرسیون تک متغیره خطی برابر یک می‌باشد.

۲-۳- خوشه‌بندی

خوشه‌بندی اشاره به روش‌های بسیار گسترده‌ای برای پیدا کردن زیرگروه یا خوشه در یک مجموعه داده دارد. وقتی مجموعه‌ای از داده‌ها خوشه‌بندی می‌شوند، تلاش می‌شود تا داده‌ها به گروه‌های متمایزی تقسیم‌بندی شوند به نحوی که داده‌های موجود در هر گروه بسیار به هم شبیه باشند و داده‌های هر گروه با دیگر گروه‌ها کاملاً متفاوت باشند. برای اینکه خوشه‌بندی به درستی انجام شود باید به وضوح تعریف شود که شباهت یا تفاوت بین دو یا چند مشاهده به چه معنی است [۳۳].

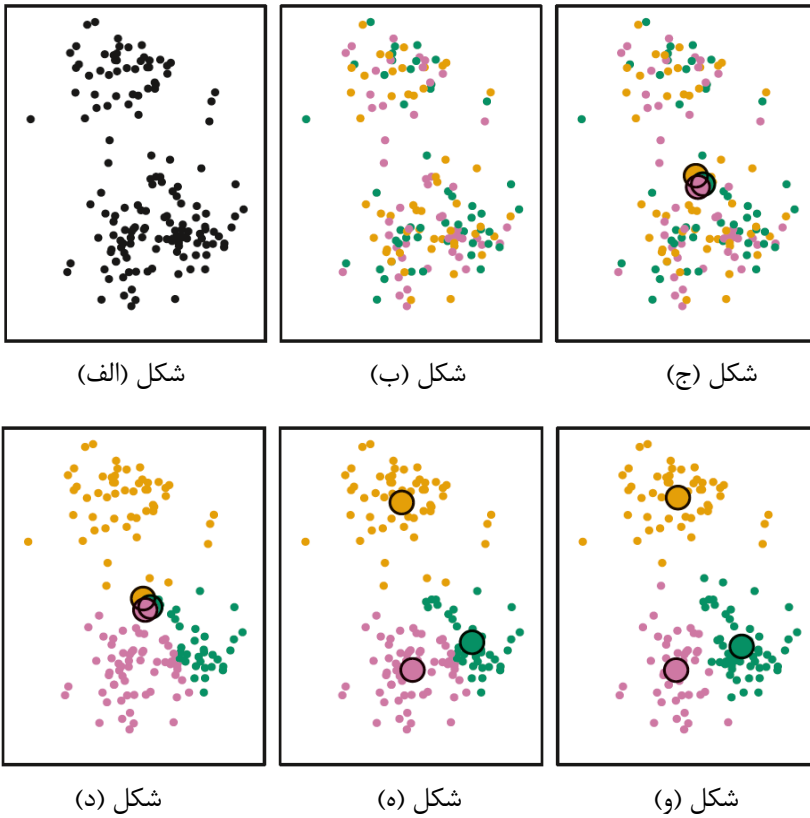
۱-۲-۳- الگوریتم خوشه‌بندی K میانگین

در خوشه‌بندی هر افزاز یک خوشه نامیده می‌شود. اعضای هر خوشه با توجه به ویژگی‌هایی که دارند به یکدیگر بسیار شبیه هستند و در عوض میزان شباهت بین خوشه‌ها کمترین مقدار است. با توجه به روش‌های مختلف اندازه‌گیری شباهت یا الگوریتم‌های تشکیل خوشه، ممکن است نتایج خوشه‌بندی برای مجموعه داده ثابت متفاوت باشند. الگوریتم K میانگین از دسته الگوریتم‌هایی است که در آن باید تعداد خوشه‌ها از قبل برای الگوریتم مشخص باشد. برای مشخص کردن شباهت داده‌ها از معیار و راه‌های مختلفی استفاده می‌شود که یکی از آن‌ها فاصله اقلیدسی است. بعد از انتخاب تعداد خوشه‌ها (K)، هر یک از نقاط به صورت تصادفی به یکی از خوشه‌ها اختصاص داده می‌شوند. سپس تا هنگامی که خوشه‌ها دیگر تغییر نکنند فرآیند زیر تکرار می‌شود:

- مرکز هر یک از خوشه‌ها با استفاده از میانگین‌گیری از بردار هر یک از نقاط آن خوشه به دست می‌آید.
- هر نقطه به خوشه‌ای اختصاص می‌یابد که به مرکز آن خوشه نزدیک‌تر است.

برای مثال در شکل ۲ (الف) داده‌ها به صورت خام و خوشه‌بندی نشده دیده می‌شوند. سپس در شکل ۲ (ب) هر یک از نقاط به صورت اتفاقی به یکی از خوشه‌ها اختصاص داده شده است. مرکز هر یک از خوشه‌ها با استفاده از میانگین‌گیری از بردار هر یک از نقاط آن خوشه به دست می‌آید (شکل ۲ (ج)). سپس هر نقطه به خوشه‌ای اختصاص می‌یابد که به مرکز آن خوشه نزدیک‌تر است (شکل ۲ (د)). مرکز هر یک از خوشه‌های جدید محاسبه می‌شود (شکل ۲ (ه)). این فرآیند تکرار می‌شود تا هیچ یک از خوشه‌ها دیگر

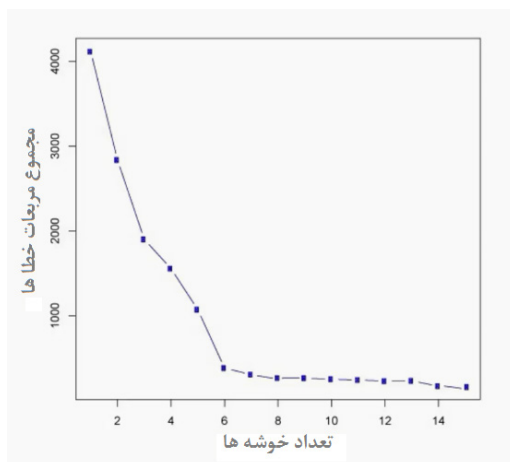
تغییر نکنند (شکل ۲ (و)) [۳۳].



شکل ۲: مراحل تشکیل خوشه‌بندی به روش K میانگین [۳۳]

حال، یک سؤال مهم مطرح می‌شود و آن اینکه چگونه تعداد خوشه‌ها K را مشخص کنیم؟ پاسخ ساده‌ای برای انتخاب بهترین تعداد خوشه‌ها یا K وجود ندارد اما یکی از راه‌هایی که می‌توان این عدد را مشخص کرد روشی به نام روش آرنج است. در این روش در ابتدا مجموع مربعات خطاها به ازای هر یک از مقادیر مختلف K محاسبه می‌شود. برای مثال $K=2, 4, 6, 8$. مجموع مربعات خطاها به معنی مجموع مربع فاصله هر یک از نقاط خوشه‌ها با مرکز آن خوشه است. اگر K را نسبت به مجموع مربعات خطاها رسم کنیم خواهیم دید که با افزایش مقدار K مقدار مجموع مربعات خطاها کاهش خواهد یافت. دلیل این امر این است که با افزایش تعداد خوشه‌ها، خوشه‌ها کوچکتر خواهند بود بنابراین انحراف نیز کوچک‌تر خواهد بود. ایده کلی روش آرنج این است که مقداری را برای K را انتخاب کنیم که به ازای آن مقدار مجموع مربعات خطاها ناگهان به مقدار زیادی کاهش یابد. برای مثال در شکل ۳ می‌توانیم ببینیم که این اتفاق در حوالی $K=6$ رخ داده است و در نمودار چیزی شبیه به آرنج ظاهر شده است. در واقع مقداری را برای

K انتخاب می‌کنیم که با افزایش K اطلاعات بیشتر چندانی به دست نمی‌آوریم یعنی دیگر با افزایش K مقدار مجموع مربعات خطاها چندان کاهش نمی‌یابد. این نکته را باید در نظر داشت که جواب صد در صد درستی برای تعداد خوشه‌ها وجود ندارد و باید با توجه به کاری که قصد انجام آن را داریم و حوزه دانش و تجربیاتمان مقدار K را انتخاب کنیم [۳۳].



شکل ۳: تشکیل آرنج بر روی منحنی [۳۳]

استفاده از الگوریتم K میانگین دارای مزیت زمانی است چراکه پیچیدگی زمانی آن، خطی می‌باشد.

۳-۳- سیستم‌های توصیه‌گر

سیستم‌های توصیه‌گر یکی از انواع سیستم‌های یادگیری ماشین هستند که با تحلیل رفتار کاربر خود، اقدام به پیشنهاد مناسب‌ترین اقلام (داده، اطلاعات، کالا و غیره) به وی می‌نماید. این نوع سیستم رویکردی است که برای مواجهه با مشکلات ناشی از حجم فراوان و رو به رشد اطلاعات ارائه شده است و به کاربر خود کمک می‌کند تا در میان حجم عظیم اطلاعات، سریع‌تر به هدف خود نزدیک شود. در پشت صحنه، سیستم‌های توصیه‌گر قدرت خودشان را از توابع توصیه‌گر می‌گیرند. توابع توصیه‌گر (شکل ۳) اطلاعات کاربر را دریافت می‌کنند و پیش‌بینی می‌کنند که کاربر به هر آیتم چه امتیازی خواهد داد، بدون اینکه کاربر حتی آیتم یا محصول را دیده باشد. این بدین معنی است که سیستم می‌تواند فقط محصولاتی را به کاربر نشان دهد که کاربر بیشتر از همه علاقه دارد و وقت کاربر را با محصولاتی که برای او مناسب نیستند تلف نکند.



شکل ۴: توابع توصیه گر [۳۶]

به عنوان مثال شما در یک سیستم توصیه گر مخصوص خرید کتاب در حال جست و جو هستید. سیستم اطلاعات خریدهای قبلی شما و امتیازات بالا یا پایینی که به کتابهای دیگر داده اید را دارد و با استفاده از این اطلاعات پیش بینی می کند شما به هر کتاب احتمالاً چه امتیازی خواهید داد. پس ابتدا کتابهایی را برای شما نمایش می دهد که شما بیشتر علاقه دارید و احتمال خرید آن توسط شما بیشتر است.

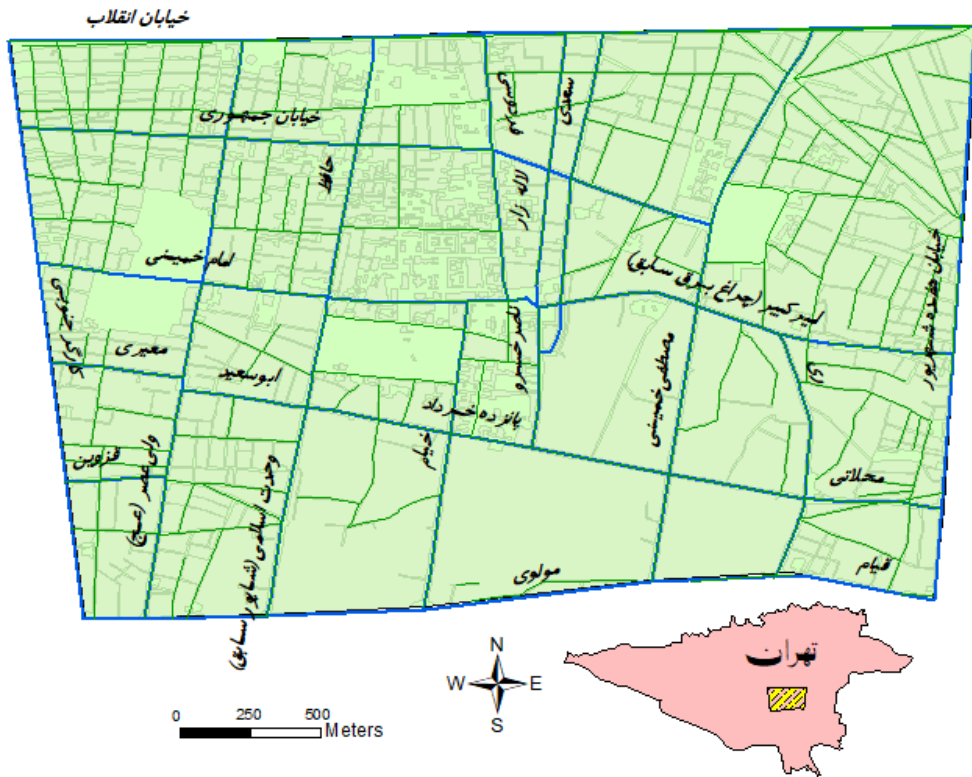
۴- روش پژوهش

در این تحقیق عوامل و پارامترهای مختلفی برای رسیدن به بهترین پیشنهاد ممکن برای کاربر در نظر گرفته شده است. برای رسیدن به بهترین پیشنهاد باید به روشی صحیح برای استفاده حداکثری و بهینه از میزان پول و زمان موجود رسید. باید زمانی را که کاربر می تواند در راه رسیدن به مقصد باشد، در نظر گرفت. اینکه فرد پیاده یا سواره است نیز عامل بسیار تأثیرگذار بر پیشنهادها است. اطلاعات مربوط به ساعت، فصل، ماه، روز، دما، شرایط آب و هوایی، میزان پول و زمان مورد نیاز برای هر مرکز و ساعات کاری آن و غیره در نظر گرفته می شود. همچنین در این تحقیق سعی شده که مسائل مختلف روانشناسی افراد در هنگام گردش و تفریح نیز مورد توجه قرار گیرد به طور مثال ممکن است مدل، مسیر دسترسی به مقصد و زمان رسیدن را تا حدی طولانی تر کند تا کاربر بتواند از مسیری به مقصد برسد که زیباتر و دلچسب تر است. شرایط آب و هوایی و ساعات کار مراکز نیز در نظر گرفته شده اند. تحقیقات نشان داده است مردم از پیشنهادهایی که بافت آگاه هستند بیشتر راضی هستند. مثلاً مردم عموماً ترجیح می دهند موزه یا بناهای تاریخی را هنگامی که خلوت هستند و در هوای آفتابی و نه بارانی، بازدید کنند و یا قدم زدن در شهر را در روز انجام دهند تا شب [۶]. گام اول در این تحقیق جمع آوری اطلاعات مربوط به شرایط و موقعیت فرد است. در ابتدا سیستم سؤالاتی از این نوع را می پرسد: میزان پولی که کاربر تمایل به خرج کردن دارد؟ کل میزان زمانی که فرد می تواند برای این گردش شهری اختصاص دهد؟ حداکثر زمانی که فرد ترجیح می دهد در راه رسیدن به مرکز صرف کند؟ آیا کاربر پیاده است یا سوار بر خودرو؟ موقعیت

فعلی کاربر در شهر کجاست؟ آیا فرد ترجیح خاصی در بازدید از نوع خاصی از مراکز مثلاً موزه‌ها یا مراکز ورزشی دارد یا خیر؟ سپس سیستم با در نظر گرفتن این اطلاعات و اطلاعات دیگری همچون بازه زمانی یا مالی مورد نیاز برای هر محل، محل‌های قابل دسترسی برای فرد، نوع مطلوب محل برای فرد، روز هفته، باز بودن یا بسته بودن محل، دمای هوا و غیره، شروع به یافتن و پیشنهاد دادن محل‌های مناسب برای فرد می‌کند و در صورتی که زمان و پولی باقی بماند، سیستم به سراغ یافتن پیشنهاد دوم برای فرد می‌رود. مثلاً سیستم با ارائه یک مسیر به فرد پیشنهاد می‌دهد که به موزه A برود و چقدر پول و زمان برای بازدید مناسب از این موزه لازم است و اگر پول و زمان کافی باقی ماند به کاربر پیشنهاد دهد که از چه مسیری می‌تواند به یک رستوران B که مناسب با زمان دسترسی فرد و پول باقی مانده است، برود. در انتها از کاربر پرسیده می‌شود که میزان رضایت خودش را با یک تا پنج ستاره، از مسیرهای پیشنهادی سیستم بدهد. در مرحله بعد سیستم با توجه به داده‌های دریافتی و سفرهایی که امتیاز دریافت کرده‌اند و تحلیل این داده‌ها با استفاده از علوم داده و یادگیری ماشین سعی در یافتن نتایج و الگوهای برای پیش‌بینی رفتار و نیازهای آینده کاربران می‌کند. مثلاً مردم چه مناطقی بیشتر تمایل به چه نوع مراکز تفریحی دارند؟ یا مردم در روزهای آخر هفته که دما بالایی داشته باشد تمایل دارند به چه محل‌هایی بروند؟

۵- پیاده‌سازی و تحلیل نتایج

برای پیاده‌سازی تحقیق، منطقه مرکزی شهر تهران که شامل بازار و بافت قدیمی اطراف می‌شود مطابق شکل ۵ در نظر گرفته شد. دلیل انتخاب این منطقه، تراکم بافت تاریخی به انضمام کافه و رستوران در آن است که از اماکن مورد علاقه به شمار می‌روند. اطلاعات پایه در مورد اماکن مورد علاقه از Google، OSM، WikiMapia و چند منبع اینترنتی دیگر اخذ گردید و با بررسی میدانی، اطلاعات مراکز مورد علاقه تا حد ممکن، تدقیق، تکمیل و بروزرسانی و برای انجام فرآیندها وارد مدل شد. در این پژوهش سامانه توسعه داده شده توسط جامعه‌ای متشکل از ۲۰ کاربر داوطلب به کار گرفته شد و نتایج کار مورد تحلیل و بررسی قرار گرفت.



شکل ۵: منطقه مورد مطالعه

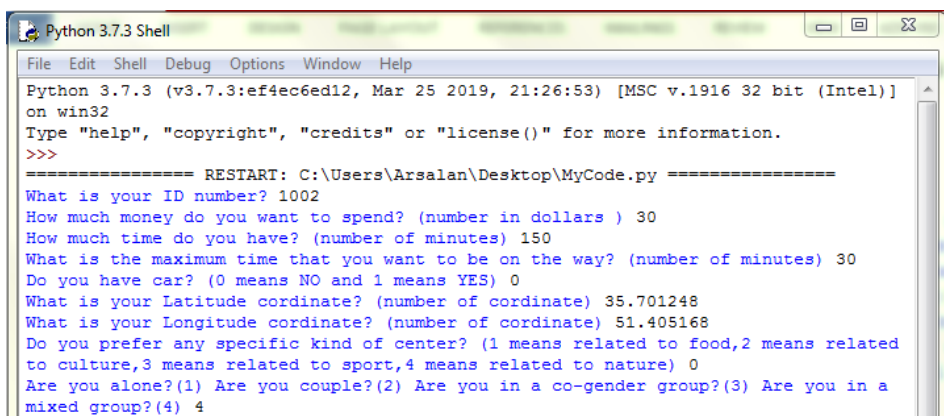
۱-۵- بخش توصیه‌گر با بهره‌گیری از علایق شخصی کاربر

برای عملکرد بهتر مدل، کاربران در هنگام ثبت نام می‌توانند اطلاعاتی از قبیل سال تولد، جنسیت، شهر محل زندگی را وارد کنند. در هر بار استفاده از سیستم، کاربر میزان پول و زمانی را که علاقه‌مند به صرف کردن برای این گردش شهری است، وارد می‌کند. همچنین از کاربر پرسیده می‌شود که موقعیت فعلی شما چیست و نهایت زمانی را که کاربر می‌خواهد در مسیر رسیدن به مقصد تفریحی باشد چه میزان است. هرچه زمانی که کاربر برای رسیدن به مقاصد تفریحی اختصاص می‌دهد بیشتر باشد مدل می‌تواند مراکز دورتری را به فهرست مراکز امکانپذیر برای گردش بیفزاید. چنانچه کاربر اعلام کند که سوار بر وسیله نقلیه است، مدل سرعت کاربر در رسیدن به مراکز را متناسب با سرعت متوسط خودروها در آن شهر یا منطقه در نظر می‌گیرد که معمولاً در شهرهای بزرگ این سرعت ۵ تا ۷ برابر سرعت پیاده‌روی معمولی فرد است. در نهایت از کاربر پرسیده می‌شود که آیا نوع خاصی از مراکز را در حال حاضر ترجیح می‌دهد یا خیر. به طور مثال فردی ترجیح می‌دهد که به گردش در محل‌های تاریخی و فرهنگی بپردازد و

فردی دیگر علاقه‌مند به صرف پول و زمان در مراکز ورزشی است. سؤالات پرسیده شده از کاربر در جدول ۱ نمایش داده شده است. در شکل ۶ سؤالات پرسیده شده از کاربر در برنامه نوشته شده برای مدل در محیط برنامه‌نویسی پایتون نمایش داده شده است.

جدول ۱: سؤالات پرسیده شده از کاربر

۱۰۰۲	لطفاً شناسه خود را وارد کنید؟
۳۰ دلار	حداکثر پولی که قصد خرج کردن دارید، چه میزان است؟
۱۵۰ دقیقه	حداکثر زمانی که می‌توانید اختصاص دهید، چه میزان است؟
۳۰ دقیقه	حداکثر زمانی که تمایل دارید در راه باشید، چقدر است؟
خیر	آیا سوار بر خودرو هستید؟
۳۵/۷۰۱۲۴۸	در چه طول جغرافیایی قرار دارید؟
۵۱/۴۰۵۱۶۸	در چه عرض جغرافیایی قرار دارید؟
خیر	آیا نوع خاصی از محل‌ها را در نظر دارید؟
گروه	آیا تنها هستید یا با گروهی از افراد هستید؟



شکل ۶: سؤالات پرسیده شده از کاربر در محیط برنامه‌نویسی پایتون

داده‌های حاصل از جواب‌های کاربر در جدولی مانند جدول ۲ ذخیره می‌شوند.

جدول ۲: داده‌های حاصل از پاسخ به سؤالاتی که از کاربر پرسیده می‌شود

	J	K	L	M	N	O	P	Q	R
ID	Money	Total Time	Travel Time	Have Car	User LAT	User LONG	Prefer	Alone or Group	
1002	30	150	30	0	35.70125	51.405168	0		4
1005	15	120	20	1	35.70397	51.399164	1		3
1003	27	100	15	0	35.73074	51.490551	2		2
1002	28	90	25	1	35.81579	51.423723	2		2
1080	60	110	40	1	35.83054	51.466416	3		2
1102	23	47	11	1	35.86758	51.481122	4		3
1006	19	58	13	1	35.90462	51.495828	0		1
1004	28	54	15	0	35.94166	51.510534	0		1
1008	20	77	17	0	35.9787	51.52524	2		2
1106	33	24	18	1	36.01574	51.539947	3		4
1452	23	78	12	0	36.05278	51.554653	4		4

در هر بار استفاده هر یک از کاربران از سیستم، اطلاعات زمینه‌ای سفر ثبت می‌شوند. این اطلاعات شامل شناسه منحصر به فرد هر کاربر، تاریخ، فصل، ماه، روز و ساعت درخواست سفر همچنین دما، وضعیت آب و هوایی و میزان آلودگی هوا در شهری که فرد درخواست مسیر گردشی کرده است. وجود این اطلاعات زمینه‌ای به عملکرد بهتر سیستم کمک می‌کند. اطلاعات زمینه‌ای هر مسیر گردشی در جدول ۳ نمایش داده شده است.

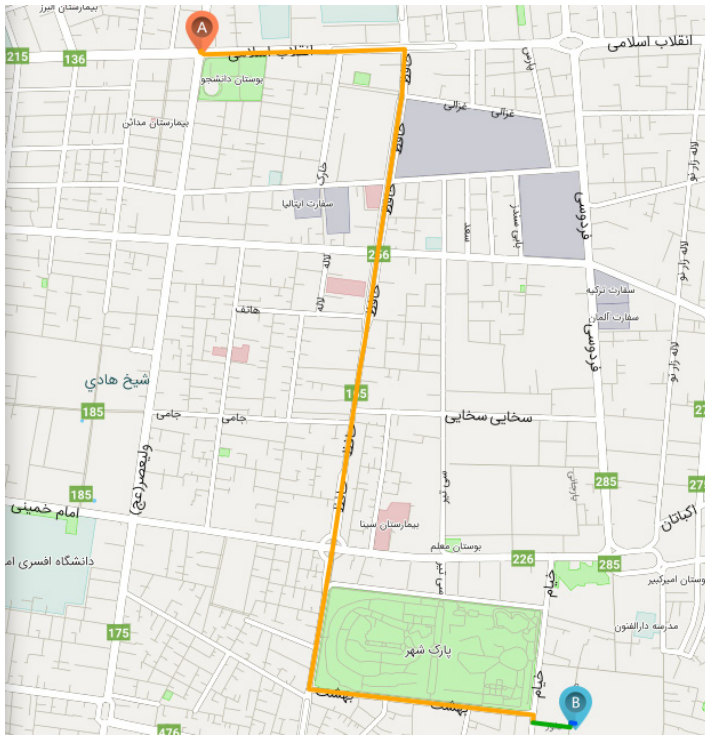
جدول ۳: داده‌های زمینه‌ای هر مسیر تفریحی

	A	B	C	D	E	F	G	H	I
1	User ID	Date	Season	Month	Day	Hour	Temperature	Weather	Pollution
2	2980002	1/1/2017	Winter	January	Sunday	2:51 AM	27	Sunny	4
3	2980003	1/2/2017	Winter	January	Monday	6:30 PM	28.9	Cloudy	2
4	2980004	1/3/2017	Winter	January	Tuesday	10:09 AM	34.5	Rainy0	9
5	2980005	1/4/2017	Winter	January	Wednesday	1:48 AM	44.1	Snowy0	1
6	2980006	1/5/2017	Winter	January	Thursday	5:27 PM	42.4	Foggy	7
7	2980007	1/6/2017	Winter	January	Friday	9:06 AM	25.3	Sunny	8
8	2980008	1/7/2017	Winter	January	Saturday	12:45 AM	32.9	Cloudy	2
9	2980009	1/8/2017	Winter	January	Sunday	4:24 PM	37.5	Rainy1	0
10	2980010	1/9/2017	Winter	January	Monday	8:03 AM	38.1	Snowy	0
11	2980011	1/10/2017	Winter	January	Tuesday	11:42 PM	43.4	Foggy	0
12	2980012	1/11/2017	Winter	January	Wednesday	3:21 PM	32.6	Sunny	6
13	2980013	1/12/2017	Winter	January	Thursday	7:00 AM	38.2	Cloudy	7
14	2980014	1/13/2017	Winter	January	Friday	10:39 PM	37.5	Rainy	1
15	2980015	1/14/2017	Winter	January	Saturday	2:18 PM	44.1	Snowy	0
16	2980016	1/15/2017	Winter	January	Sunday	5:57 AM	43.4	Fog	1
17	2980017	1/16/2017	Winter	January	Monday	9:36 PM	30.6	Sunny	2
18	2980018	1/17/2017	Winter	January	Tuesday	1:15 PM	32.2	Cloudy	3
19	2980019	1/18/2017	Winter	January	Wednesday	4:54 AM	42.8	Rainy	8
20	2980020	1/19/2017	Winter	January	Thursday	8:33 PM	43.1	Snowy	1
21	2980021	1/20/2017	Winter	January	Friday	12:12 PM	31.6	Fog	1
22	2980022	1/21/2017	Winter	January	Saturday	3:51 AM	36.2	Sunny	6
23	2980023	1/22/2017	Winter	January	Sunday	7:30 PM	40.8	Cloudy	5

پس از پرسیدن مجموعه سؤالات مذکور و در نظر گرفتن پارامترهای زمینه‌ای، مدل شروع به پیدا کردن مراکز امکان‌پذیر برای کاربر می‌کند. باز و بسته بودن هر یک از مراکز و همچنین اینکه نیاز به رزرو دارند یا خیر، در محیط روباز یا سرپوشیده هستند یا خیر و شرایط آب و هوایی نیز مد نظر قرار می‌گیرد به گونه‌ای که اگر آن مرکز بسته باشد و یا هوا بسیار سرد یا آلوده باشد و مرکز سرپوشیده نباشد، حذف می‌شود. سپس مراکزی که متناسب با پول و زمان یا موقعیت فعلی کاربر نباشد حذف می‌شوند. برای بررسی اینکه کاربر با توجه به حداکثر زمان اختصاص داده شده برای رسیدن به مقصد و موقعیت فعلی‌اش، توانایی دسترسی به چه محل‌هایی را دارد، نیاز به استفاده و تعریف مفاهیمی همچون قاب دسترسی و فاصله منتهن است. برای بررسی این‌که کاربر پول و زمان کافی برای استفاده از یک محل را دارد یا خیر برای هر یک از مراکز تفریحی یا فرهنگی حداقل و حداکثر پول و حداقل و حداکثر زمان مورد نیاز تعریف می‌شود. با توجه به اهمیت مسیر پیشنهادی برای رسیدن به مقصد برای کاربر، مفهومی به نام نقاط طعم‌دهنده تعریف شده است تا مسیری به کاربر پیشنهاد شود که بیشترین جاذبه‌های گردشگری را دارد. مسیرهای مختلفی با توجه به موقعیت نقاط طعم‌دهنده انتخاب می‌شود. در نهایت پس از در نظر گرفتن شروط مختلف از جمله محدودیت‌های مکانی، مالی، دسترسی و ترجیح کاربر، مراکز امکان‌پذیر و مناسب برای کاربر بر روی نقشه به نمایش در می‌آیند و همچنین متنی برای کاربر ظاهر می‌شود که توضیح می‌دهد فاصله کاربر تا محل‌های مورد نظر چقدر است و چگونه می‌تواند این محل‌ها را بیابد و همچنین زمان مورد نیاز هر کدام از این محل‌ها به چه میزان است و کاربر چند دقیقه در راه خواهد بود تا به مقصد برسد. برای مثال به یک کاربر پیاده که در چهارراه ولی عصر قرار دارد و حدود ۴ ساعت زمان و تا ۵۰ دلار پول برای خرج کردن دارد و علایق او کافه‌گردی، بازی و بازدید از اماکن فرهنگی است، پیشنهادهای زیر از طریق برنامه ارائه شده است:

- رستوران ولیمه در ۴۰ متر به سمت چپ و ۴۸۰ متر به سمت پایین قرار دارد و این مکان نیاز به ۵ تا ۱۵ دلار و ۴۰ تا ۶۰ دقیقه زمان دارد و شما ۵ دقیقه در راه خواهید بود.
 - کاخ گلستان در ۸۹۰ متر به سمت راست و ۱۰۴۲ متر به سمت پایین قرار دارد و این مکان نیاز به ۲۰ تا ۴۰ دلار و ۹۰ تا ۲۴۰ دقیقه زمان دارد و شما ۲۳ دقیقه در راه خواهید بود.
 - کافه سنتی خانه هنر میرا در ۳۰۰ متر به سمت چپ و ۲۸۰ متر به سمت پایین قرار دارد و این مکان نیاز به ۱۰ تا ۲۰ دلار و ۳۰ تا ۶۰ دقیقه زمان دارد و شما ۶ دقیقه در راه خواهید بود.
- شایان ذکر است که برنامه توسعه داده شده در حال حاضر به صورت رومیزی است ولی نتایج روی نقشه

نشان داده می‌شود و مسیر نیز نوشته می‌شود. شکل ۷ نمایش مسیر همان گردشگر از چهارراه ولیعصر تا کاخ گلستان را بر روی سامانه نقشه تهران نشان می‌دهد.



شکل ۷: کاخ گلستان، مقصد پیشنهاد شده به گردشگری که در چهارراه ولی عصر قرار دارد

۲-۵- بخش توصیه‌گر بر اساس سوابق کاربران

با استفاده کاربران از سیستم، اطلاعات جمع‌آوری شده، منبع خوبی برای بهبود عملکرد و ارائه پیشنهادهای بهتر و دقیق‌تر خواهد بود. بنابراین در پیشنهادهای بعدی به کاربران می‌توان سوابق را در نظر گرفت که این کار با بهره‌گیری از ماشین‌های بردار پشتیبان به عنوان یک الگوریتم مخصوص طبقه‌بندی انجام شده است. در اینجا پیش‌بینی می‌شود که دما، رطوبت، شرایط جوی و آب و هوایی برای ورزش و گردش در فضای باز مناسب است یا خیر. نحوه کار به این صورت است که با وارد کردن مقادیر این متغیرها به عنوان داده ورودی، می‌توان جوابی به صورت مثبت یا منفی از الگوریتم دریافت کرد که نشان دهنده مساعد بودن شرایط جوی برای ورزش و تفریح در فضای باز است (جدول ۴). مدل توصیه‌گر می‌تواند با در نظر گرفتن این شرایط زمینه‌ای و با استفاده از الگوریتم طبقه‌بندی ماشین‌های بردار پشتیبان، در شرایطی که اوضاع

برای حضور در فضای باز مساعد نیست، تنها به پیشنهادهای مکان‌های سر پوشیده بپردازد.

جدول ۴: شرایط توصیه شده بر اساس ماشین‌های بردار پشتیبان

	A	B	C	D	E
1	Temperature	Outlook	Humidity	Windy	OK?
2	Mild	Sunny	80	No	Yes
3	Hot	Sunny	75	Yes	No
4	Hot	Overcast	77	No	Yes
5	Cool	Rain	70	No	Yes
6	Cool	Overcast	72	Yes	Yes
7	Mild	Sunny	77	No	No
8	Cool	Rain	70	No	Yes
9	Mild	Sunny	69	No	Yes
10	Mild	Overcast	65	Yes	No
11	Mild	Overcast	73	Yes	No

۳-۵- استفاده از خوشه‌بندی

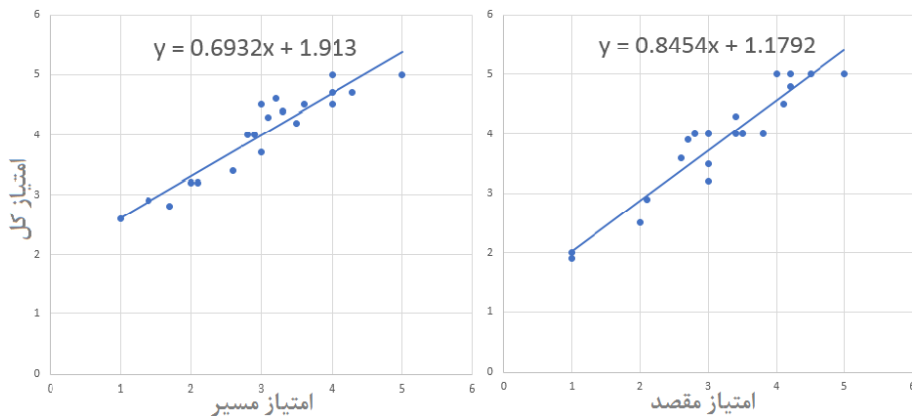
در این قسمت با استفاده از خوشه‌بندی کاربران و مراکز سعی شده است تا مناسب‌ترین گزینه برای هر کاربر با توجه به انتخاب‌های قبلی او پیشنهاد شود. در مرحله نخست، ابتدا متناسب با انواع مختلف نقاط مورد علاقه و کاربری‌ها، میانگین پول و زمانی که هر کاربر در طول تمامی استفاده‌هایش از سیستم صرف کرده است، محاسبه می‌شود. سپس این میانگین‌ها برای تمامی کاربران در یک نمودار هزینه و زمان رسم می‌شوند. بعد از آن، نقاط این نمودار با استفاده از الگوریتم K میانگین خوشه‌بندی می‌شوند. تعداد خوشه‌های به دست آمده متناسب با داده‌های ورودی می‌تواند متفاوت باشد. فرض کنیم که الگوریتم K میانگین کاربران امکانات ورزشی را به چهار خوشه متفاوت افراز کند. خوشه اول کاربرانی که پول و زمان زیادی را در مراکز ورزشی خرج می‌نمایند. خوشه دوم کاربرانی که پول زیاد اما زمان کمی را در مراکز ورزشی صرف می‌نمایند. خوشه سوم کاربرانی که پول کم اما زمان زیادی را در مراکز ورزشی خرج می‌کنند و در نهایت خوشه چهارم، کاربرانی که پول و زمان کمی را در مراکز ورزشی صرف می‌کنند.

در مرحله دوم، انواع مختلف نقاط مورد علاقه خوشه‌بندی می‌شوند. به این صورت که ابتدا میانگین حداقل زمان مورد نیاز و حداکثر زمان مورد نیاز برای استفاده از مراکز مختلف، مثلاً ورزشی محاسبه می‌شود. سپس از حداقل و حداکثر هزینه مورد نیاز برای هر یک از مراکز ورزشی میانگین‌گیری شده، این مراکز با توجه به میانگین پول و زمان محاسبه شده، در یک نمودار هزینه و زمان رسم می‌شوند. سپس، برابر با

تعداد خوشه‌های یافت شده برای کاربران در مرحله اول، مراکز ورزشی خوشه‌بندی می‌شوند. پس مراکز ورزشی نیز به چهار خوشه متفاوت متناسب با هزینه و زمان مورد نیاز، افزاز شده‌اند. در مرحله سوم، متناسب با موقعیت فرد و شرایط وی چنانچه مراکز ورزشی از انواع مختلف خوشه‌بندی برای کاربر امکان‌پذیر باشد، مرکزی که متناسب با خوشه‌بندی کاربر باشد به عنوان پیشنهاد اول نمایش داده می‌شود و سپس به ترتیب مراکز خوشه‌های متفاوت‌تر و دورتر. این نوع خوشه‌بندی برای انواع مختلف مراکز مانند تفریحی، فرهنگی، رستوران‌ها و غیره نیز انجام می‌شود. به عنوان مثال چنانچه فرد امکان استفاده از یک رستوران از خوشه اول و امکان استفاده از رستوران دیگری از خوشه چهارم را داشته باشد، رستوران متعلق به خوشه اول به عنوان پیشنهاد نخست ارائه می‌شود چون متناسب با انتخاب‌های قبلی او، این کاربر متمایل به استفاده از رستوران‌های لوکس بوده است.

۴-۵- رگرسیون امتیازها

در این قسمت پیش‌بینی میزان رضایت کلی کاربر نسبت به رضایت از محل پیشنهادی و همچنین میزان رضایت کلی کاربر نسبت به میزان رضایت کاربر نسبت به مسیر پیشنهادی بررسی می‌شود. هر کاربر پس از استفاده از مقصد توصیه شده، میزان رضایت خود را از آن مقصد و مسیر و رضایت کلی خود را نیز اعلام می‌نماید. شکل ۸ نمودار پراکندگی امتیاز مسیر و امتیاز کل و همچنین نمودار پراکندگی امتیاز مقصد و امتیاز کل را نمایش داده است.



شکل ۸: نمودار پراکندگی امتیاز مسیر و امتیاز کل و همچنین نمودار پراکندگی امتیاز مقصد و امتیاز کل

با استفاده از رگرسیون می‌توان روند آینده و همچنین عوامل اصلی تأثیرگذار در عملکرد رضایت بخش سیستم را یافت. میزان رضایت از محل یا مسیر پیشنهادی، متغیر مستقل و میزان رضایت کلی کاربر از برنامه تفریحی متغیر وابسته است. نمودار پراکندگی امتیاز مسیر و امتیاز کل و همچنین نمودار پراکندگی امتیاز مقصد و امتیاز کل به ترتیب در نمودارهای شکل ۸ آورده شده‌اند. شیب‌های این خطوط ضریب رگرسیون هستند که این مقدار برای نمودار سمت راست در شکل ۸، $0/84$ است که بدین معنی است که چنانچه بتوانیم امتیاز رضایت کاربران از مقاصد پیشنهادی را یک واحد افزایش دهیم، رضایت کلی کاربر از برنامه تفریحی پیشنهادی به میزان $0/84$ افزایش می‌یابد. با توجه به درجه آزادی ۱۸ و بازه اطمینان ۹۵ درصد، مقدار t برابر با $2/101$ است، بنابراین مقدار بازه اطمینان برای ضریب رگرسیون برابر با $0/113$ است و به این معنی است که بیشتر از همه احتمال دارد که مقدار واقعی ضریب رگرسیون برابر با $0/84$ باشد و به احتمال ۹۵ درصد مقدار واقعی آن بین $0/71$ تا $0/97$ است. بازه اطمینان برای رگرسیون برای نمودار سمت چپ در شکل ۸ برابر با $0/12$ است و مقدار واقعی ضریب رگرسیون برای این نمودار به احتمال ۹۵ درصد بین $0/57$ تا $0/81$ است.

با بررسی این دو بازه و ضریب رگرسیون می‌توان گفت که به احتمال بیشتر، رضایت کاربر از مقصد نقش مهم‌تری از رضایت کاربر از مسیر، در رضایت کلی کاربر از برنامه پیشنهادی دارد. اما به دلیل وجود همپوشانی بین بازه‌های اطمینان این دو عامل، نمی‌توان این نتیجه‌گیری را با قطعیت کرد و چنانچه نمونه‌گیری دیگری انجام شود و بررسی تکرار شود، ممکن است این مقادیر متفاوت باشد. هر چه تعداد نمونه‌گیری‌ها بیشتر باشد، بازه اطمینان کوچکتر خواهد بود چون داده‌های بیشتری برای بررسی وجود دارد و پیش بینی دقیق‌تر خواهد بود. چنانچه امتیاز کاربر به یک مسیر یا محل را داشته باشیم حتی اگر این امتیازها از یک مسیر تفریحی دیگر به دست آمده باشد، می‌توان امتیاز کلی کاربر را به مسیر پیشنهادی جدید به دست آورد.

۶- نتیجه‌گیری و پیشنهادها

در این پژوهش، هدف طراحی سیستمی بوده است که بتواند بهترین پیشنهاد را تا حد ممکن در شرایط مختلف و برای کاربران مختلف ارائه دهد. ویژگی سیستم طراحی شده این است که برای سفرهای کوتاه با مدت زمان چندساعت مناسب می‌باشد و قادر است یک یا دو پیشنهاد را برای استفاده از وقت به کاربر ارائه دهد. این شرایط در بسیاری از سفرهای کاری رخ می‌دهد. سیستم در مواردی که هیچگونه اطلاعاتی درباره کاربر و امتیازدهی‌های وی به مکان‌های مختلف یا مسیرهای تفریحی مختلف ندارد یا کاربر در یک

شهر جدید قرار دارد که با آن به کلی ناآشناست، از روش‌های سیستم‌های توصیه‌گر بر مبنای محبوبیت استفاده می‌کند و محل‌هایی با بیشترین تعداد رأی‌دهی را یا توجه به نیاز کاربر به وی پیشنهاد می‌دهد. این روش در صورتی مناسب خواهد بود که اطلاعات شخصی و سلیق کاربر به صورت کامل و به درستی وارد سیستم شده باشند.

با داشتن اطلاعات چند مسیر تفریحی طی شده کاربر و امتیاز کاربر به آن‌ها می‌توان پروفایل شخصی آن کاربر را تشکیل داد و از آن برای دادن پیشنهادهای بعدی استفاده کرد به ویژه هنگامی که همان کاربر در منطقه یا شهری جدید قرار گرفته است یا اساساً اطلاعاتی درباره امتیازدهی به مراکز مختلف آن شهر در سیستم ما وجود ندارد و به نوعی سیستم برای اولین بار در حال پیاده شدن در یک شهر جدید است. برای انجام این کار نوعی جدید از سیستم‌های توصیه‌گر محتوا-محور در این تحقیق، ارائه شد. ضعف این نوع سیستم در حالتی خواهد بود که اطلاعات پروفایل کاربر با آنچه در ذهن او می‌گذرد منطبق نباشد. این نوع مشکلات با استفاده از سایر سیستم‌های توصیه‌گر از جمله سیستم‌های توصیه‌گر بر مبنای مشارکت حل می‌شوند. با بررسی داده‌های مربوط به پروفایل کاربری کاربران در یک شهر می‌توان به اطلاعاتی از قبیل اینکه مردم در آن شهر یا منطقه بیشتر تمایل به انجام چه نوع تفریح یا فعالیت‌هایی هستند، دست یافت ولی در این حالت احتیاج به مشارکت طیف وسیعی از مردم وجود دارد و یک سیستم باید به حد کافی بزرگ باشد تا بتواند از حالت مشارکتی استفاده نماید.

در این تحقیق به پیش‌بینی میزان رضایت کلی کاربر نسبت به رضایت از مقصد پیشنهادی و نیز نسبت به مسیر پیشنهادی پرداخته شد. با در نظر گرفتن داده‌های مربوط به امتیاز مسیر، امتیاز مقاصد و امتیاز کل و استفاده از تحلیل رگرسیون خطی میزان تأثیر هر یک از این عوامل بر رضایت کاربر مشخص گردید تا در پیشنهادهای بعدی، سیستم اصلاح گردد. نتایج به دست آمده نشان داد که به احتمال بیشتر، رضایت کاربر از مقصد نقش مهم‌تری از رضایت کاربر از مسیر، در رضایت کلی کاربر از برنامه پیشنهادی دارد. نسب به دست آمده بین این دو رضایتمندی ۸۴ به ۶۹ یعنی حدود ۱/۲ است ولی تنها در زمانی می‌توان به ضرس قاطع بر آن پای فشرده که تعداد کاربران سیستم به عدد قابل توجهی برسد. در حال حاضر، سیستم توصیه‌گر توسعه داده شده به صورت رومیزی توسط تعداد ۲۰ کاربر داوطلب مورد استفاده قرار گرفت. در توسعه‌های آتی نیاز است که این سامانه به صورت تحت وب و نیز بر روی تلفن همراه قابل استفاده گردد تا بتواند مورد استفاده کاربران زیادی قرار گیرد و تمام مزایای خود را بروز دهد. همچنین زمان پاسخگویی سیستم نیز بسیار مهم خواهد بود و با افزایش تعداد کاربران باید این زمان در حد معقولی

منابع

- [1] M. Nilashi, O. Ibrahim, and K. Bagherifard, "A recommender system based on collaborative filtering using ontology and dimensionality reduction techniques," *Expert Systems with Applications*, vol. 92, pp. 507-520, 2018, doi: <https://doi.org/10.1016/j.eswa.2017.09.058>.
- [2] K. Kabassi, "Personalizing recommendations for tourists," *Telematics and Informatics*, vol. 27, no. 1, pp. 51-66, 2010, doi: <https://doi.org/10.1016/j.tele.2009.05.003>.
- [3] H. Werthner and F. Ricci, "E-commerce and tourism," *Communications of the ACM*, vol. 47, no. 12, pp. 101-105, 2004, doi: [10.1145/1035134.1035141](https://doi.org/10.1145/1035134.1035141).
- [4] R. Goyal and S. Jain Goyal, "Recommender system: An analytical report on decision making for large scale online social networks," *Materials Today: Proceedings*, 2021, doi: <https://doi.org/10.1016/j.matpr.2021.06.311>.
- [5] L. Huang, M. Fu, F. Li, H. Qu, Y. Liu, and W. Chen, "A deep reinforcement learning based long-term recommender system," *Knowledge-Based Systems*, vol. 213, p. 106706, 2021, doi: <https://doi.org/10.1016/j.knsys.2020.106706>.
- [6] D. Gavalas, C. Konstantopoulos, K. Mastakas, and G. Pantziou, "Mobile recommender systems in tourism," *Journal of Network and Computer Applications*, vol. 39, pp. 319-333, 2014, doi: <https://doi.org/10.1016/j.jnca.2013.04.006>.
- [7] F. Ricci, L. Rokach, and B. Shapira, "Introduction to Recommender Systems Handbook," in *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira, and P. Kantor Eds. Boston, MA: Springer, 2011.
- [8] I. Mazej and E. Shmueli, "A personal data store approach for recommender systems: enhancing privacy without sacrificing accuracy," *Expert Systems with Applications*, vol. 139, p. 112858, 2020, doi: <https://doi.org/10.1016/j.eswa.2019.112858>.
- [9] A. Gharahighehi, C. Vens, and K. Pliakos, "Fair multi-stakeholder news recommender system with hypergraph ranking," *Information Processing & Management*, vol. 58, no. 5, p. 102663, 2021, doi: <https://doi.org/10.1016/j.ipm.2021.102663>.
- [10] J. Lu, D. Wu, M. Mao, W. Wang, and G. Zhang, "Recommender system application developments: A survey," *Decision Support Systems*, vol. 74, pp. 12-32, 2015, doi: <https://doi.org/10.1016/j.dss.2015.03.008>.
- [11] R. D. Burke, J. H. Kristian, and C. Y. Benjamin, "Knowledge-based navigation of complex information spaces," in *Thirteenth national conference on Artificial intelligence*, 1996, vol. 1, pp. 462-468.
- [12] R. Burke, "Hybrid Recommender Systems: Survey and Experiments," *User Modeling and User-Adapted Interaction*, vol. 12, no. 4, pp. 331-370, 2002/11/01 2002, doi: [10.1023/A:1021240730564](https://doi.org/10.1023/A:1021240730564).
- [13] T. Hung-Wen and S. Von-Wun, "A personalized restaurant recommender agent for mobile e-service," in *IEEE International Conference on e-Technology, e-Commerce and e-Service*, 2004. *EEE '04*. 2004, 28-31

March 2004 2004, pp. 259-262, doi: 10.1109/EEE.2004.1287319.

[14] A. Pashtan, R. Blattler, A. Heusser, P. Scheuermann, and B. Grove, "CATIS: A Context-Aware Tourist Information System," in 4th international workshop of mobile computing, 2003.

[15] L. Martinez, R. M. Rodriguez, and M. Espinilla, "REJA: A Georeferenced Hybrid Recommender System for Restaurants," in 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, 15-18 Sept. 2009 2009, vol. 3, pp. 187-190, doi: 10.1109/WI-IAT.2009.259.

[16] J. P. Lucas, N. Luz, M. N. Moreno, R. Anacleto, A. Almeida Figueiredo, and C. Martins, "A hybrid recommendation approach for a tourism system," *Expert Systems with Applications*, vol. 40, no. 9, pp. 3532-3550, 2013, doi: <https://doi.org/10.1016/j.eswa.2012.12.061>.

[17] A. Moreno, A. Valls, D. Isern, L. Marin, and J. Borràs, "SigTur/E-Destination: Ontology-based personalized recommendation of Tourism and Leisure Activities," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 1, pp. 633-651, 2013, doi: <https://doi.org/10.1016/j.engappai.2012.02.014>.

[18] T. Ruotsalo et al., "SMARTMUSEUM: A mobile recommender system for the Web of Data," *Journal of Web Semantics*, vol. 20, pp. 50-67, 2013, doi: <https://doi.org/10.1016/j.websem.2013.03.001>.

[19] W.-S. Yang and S.-Y. Hwang, "iTravel: A recommender system in mobile peer-to-peer environment," *Journal of Systems and Software*, vol. 86, no. 1, pp. 12-20, 2013, doi: <https://doi.org/10.1016/j.jss.2012.06.041>.

[20] P. Avesani, P. Massa, and R. Tiella, "Moleskiing.it: a Trust-aware Recommender System for Ski Mountaineering," *International Journal for Infonomics*, vol. 20, no. 35, pp. 1-10, 2005.

[21] D. Fesenmaier, F. Ricci, E. Schaumlechner, K. Wöber, and C. Zanella, "DIETORECS: Travel Advisory for Multiple Decision Styles," in ENTER 2003, 2003.

[22] L. Console, I. Torre, I. Lombardi, S. Gioria, and V. Surano, "Personalized and Adaptive Services on Board a Car: An Application for Tourist Information," *Journal of Intelligent Information Systems*, vol. 21, no. 3, pp. 249-284, 2003, doi: 10.1023/a:1025506816422.

[23] A. García-Crespo, J. Chamizo, I. Rivera, M. Mencke, R. Colomo-Palacios, and J. M. Gómez-Berbis, "SPETA: Social pervasive e-Tourism advisor," *Telematics and Informatics*, vol. 26, no. 3, pp. 306-315, 2009, doi: <https://doi.org/10.1016/j.tele.2008.11.008>.

[24] S. Schiaffino and A. Amandi, "Building an expert travel agent as a software agent," *Expert Systems with Applications*, vol. 36, no. 2, Part 1, pp. 1291-1299, 2009, doi: <https://doi.org/10.1016/j.eswa.2007.11.032>.

[25] Y. Huang and L. Bian, "A Bayesian network and analytic hierarchy process based personalized recommendations for tourist attractions over the Internet," *Expert Systems with Applications*, vol. 36, no. 1, pp. 933-943, 2009, doi: <https://doi.org/10.1016/j.eswa.2007.10.019>.

[26] N. Saiph Savage, M. Baranski, N. Elva Chavez, and H. T., "I'm feeling LoCo: A Location Based Context Aware Recommendation System," in *Advances in Location-Based Services. Lecture Notes in Geoinformation and Cartography*, G. Gartner and F. Ortog Eds. Berlin, Heidelberg: Springer, 2012.

[27] J. Koza, F. H. Bennett, D. Andre, and M. Keane, "Automated Design of Both the Topology and Sizing of

- Analog Electrical Circuits Using Genetic Programming," in *Artificial Intelligence in Design* 96, 1996: Springer, pp. 151-170.
- [28] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer, 2006.
- [29] J. Friedman, "Data Mining and Statistics: What'S the Connection," *Computing Science and Statistics*, vol. 29, no. 1, pp. 3-9, 1997.
- [30] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. CreateSpace Independent Publishing Platform, 2016.
- [31] G. E. Hinton and T. J. Sejnowski, *Unsupervised Learning: Foundations of Neural Computation*. MIT Press, 1999.
- [32] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys*, vol. 41, no. 3, p. Article 15, 2009, doi: 10.1145/1541880.1541882.
- [33] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: with Applications in R*. Springer Publishing Company, Incorporated, 2014.
- [34] T. Hastie, R. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction (Springer series in statistics)*. Springer, 2009.
- [35] H. L. Seal, "Studies in the History of Probability and Statistics. XV: The Historical Development of the Gauss Linear Model," *Biometrika*, vol. 54, no. 1/2, pp. 1-24, 1967, doi: 10.2307/2333849.
- [36] T. Q. Lee, Y. Park, and Y.-T. Park, "A time-based approach to effective recommender systems using implicit feedback," *Expert Systems with Applications*, vol. 34, no. 4, pp. 3055-3062, 2008, doi: <https://doi.org/10.1016/j.eswa.2007.06.031>.