

شخصی سازی یک سیستم توصیه گر گردشگری بر مبنای تشابه کاربران و استفاده از شبکه باور عمیق

زینب نیسانی سامانی^۱، علی اصغر آل شیخ^{۲*}، ابوالقاسم صادقی نیارکی^۳، مهدی نظری اشنی^۴

۱- دانشجوی دکتری، گروه سیستم‌های اطلاعات مکانی، دانشکده مهندسی نقشه‌برداری، دانشگاه صنعتی خواجه نصیرالدین طوسی

۲- استاد گروه سیستم‌های اطلاعات مکانی، دانشکده مهندسی نقشه‌برداری، دانشگاه صنعتی خواجه نصیرالدین طوسی

۳- استادیار گروه سیستم‌های اطلاعات مکانی، دانشکده مهندسی نقشه‌برداری، دانشگاه صنعتی خواجه نصیرالدین طوسی

۴- دانشجوی کارشناسی ارشد، گروه سیستم‌های اطلاعات مکانی، دانشکده مهندسی نقشه‌برداری، دانشگاه صنعتی خواجه نصیرالدین طوسی

تاریخ دریافت مقاله: ۱۴۰۰/۰۱/۳۰ تاریخ پذیرش مقاله: ۱۴۰۰/۰۶/۲۸

چکیده

سیستم‌های توصیه گر مکانی امکان ارائه اطلاعات مفید را برای کاربران با کاهش اطلاعات تکراری و نامربوط در وب به‌طور گسترده فراهم می‌سازند. در عصر حاضر با وجود سیستم‌های توصیه گر یافتن مکان مورد علاقه در کوتاه‌ترین زمان، با صرف کم‌ترین هزینه و توسط شخص کاربر صورت می‌پذیرد. سیستم‌های توصیه گر در حوزه‌های مختلف از جمله گردشگری کاربرد زیادی دارند و به‌عنوان ابزار توسط گردشگر قابل استفاده می‌باشند. عموماً سیستم‌های توصیه گر بر پایه پالایش مشارکتی و بر اساس مشابهت میان گروه کاربران، پیشنهادات لازم را به کاربران مختلف ارائه می‌دهند. در صورتیکه بسیاری از خدمات با ویژگی‌های شخصی و خصوصیات فرد منطبق نیست و این امر از کارایی اینگونه سیستم‌ها می‌کاهد. هدف این پژوهش توسعه یک الگوریتم توصیه گر مبتنی بر شباهت بین کاربران و مفاهیم شخصی سازی است. نوآوری این تحقیق بکارگیری شبکه عصبی باور عمیق در راستای شخصی سازی پیشنهادات قابل ارائه به گردشگران است. سناریوی تحقیق به این صورت است که ابتدا گردشگران مختلف در سیستم ثبت نام می‌کنند، سپس اطلاعات فردی، ارجحیت‌های عمومی و عوامل خاص شخصی سازی خود را برای بازدید از مراکز گردشگری بیان می‌کنند. در رویکرد پیشنهادی نیازی به تفکیک کاربران نیست بلکه با توجه به قدرت یادگیری شبکه‌های عمیق، امکان تفکیک و شخصی سازی پیشنهادات کاربران وجود دارد. در این راستا داده‌های مربوط به ۴۰۰ مسافر گردشگری براساس ۱۴ المان ورودی و متمایز کننده دریافت شد. همچنین براساس شبکه آموزش دیده، قابلیت پیش‌بینی مکان‌های گردشگری شخصی سازی شده برای ۳۰ کاربر جدید بررسی شد. ارزیابی نتایج براساس سه شاخص معیار صحت، پوشش و معیار F و همچنین رضایت‌مندی کاربران انجام شد. نتایج حاصل، دقت بالا و همچنین رضایت‌مندی بیش از ۷۹ درصدی کاربران را نشان می‌دهد.

کلید واژه‌ها: سیستم توصیه گر، گردشگری، شخصی سازی، شبکه باور عمیق.

* نویسنده مکاتبه کننده: تهران- خیابان ولیعصر- تقاطع میر داماد غربی- دانشکده مهندسی نقشه‌برداری دانشگاه صنعتی خواجه نصیرالدین طوسی

۱- مقدمه

سیستم‌های توصیه‌گر گردشگری، ابزارهایی هستند که توسط آنها یک فرد علاقه‌مند به گردشگری بتواند در کوتاه‌ترین زمان ممکن، با حداقل امکانات، کم‌ترین هزینه و با دانش اندک، ظرفیت‌های گردشگری یک منطقه را شناسایی کند و با یک برنامه هدف‌دار از سوی دولت، آژانس‌های گردشگری و غیره به مکانی خاص سفر کند. هدف استفاده از سیستم‌های توصیه‌گر در زمینه گردشگری، سفارشی‌کردن و شخصی‌سازی اطلاعات موجود درباره جاذبه‌های گردشگری یک منطقه براساس ترجیحات و علایق گردشگران می‌باشد/۱. وجود یک سیستم توصیه‌گر مکانی از این نظر که می‌تواند در قالب یک تور گردشگری به کاربر یا گروهی از کاربران مکان‌های گردشگری در یک منطقه جغرافیایی را طبق اطلاعاتشان اعلام کند و همچنین طرح روزانه سفرشان را محاسبه کند، مفید می‌باشد. چنین سیستم‌هایی باعث پیشرفت صنعت گردشگری و آسایش گردشگران می‌شود/۲، ۳ و ۴. رویکردهای مبتنی بر مشابهت‌یابی کاربران به خصوص روش‌های پالایش مشارکتی (CF)^۱ از رویکردهای آماری برای تخمین شباهت‌ها استفاده می‌کنند. روش‌های پالایش مشارکتی به‌طور کلی از ارزیابی جمعی برای توصیه به کاربران استفاده می‌کند. بنابراین لازم است محققان ابتدا یک جامعه را با علایق مشترک و خصوصیات مشابه شناسایی کنند و سپس پیشنهادات را ارائه نمایند. این روش می‌تواند یک انتخاب مناسب برای توصیه به گردشگر باشد. سیستم‌های پیشنهادی مبتنی بر پالایش مشارکتی مواردی هستند که در آنها توصیه‌ها فقط شباهت بین کاربران را در نظر می‌گیرند/۵، ۶ و ۷.

در این پژوهش، دو گروه از عوامل توصیه‌گری شامل اطلاعات فردی و همچنین ارجحیت‌های شخصی شامل

۱۴ المان جنسیت، سن، شغل، درآمد ماهیانه، وضعیت تاهل، محل سکونت، توانایی جسمانی، رشته ورزشی تخصصی، علاقه به گردشگری، وسیله مورد استفاده، توانایی پیاده‌روی، توانایی کوهنوردی، نزدیکی به وسایل نقلیه عمومی و محدودیت زمانی (مدت اقامت) در نظر گرفته شد / ۷ و ۸. وظیفه سیستم توصیه‌گر کاهش زمان و هزینه در جستجوی اطلاعات از طریق ارائه پیشنهادات متناسب با نیاز و خواست کاربران است. انجام این وظیفه در مرحله اول نیاز به شناخت کاربر دارد چراکه سیستم باید اطلاعات موردنظر و دلخواه کاربر را از میان حجم عظیم اطلاعات موجود در وب تفکیک کند، شباهت میان کاربران و اقلام را بررسی نموده و مناسب‌ترین آنها را پیشنهاد دهد / ۶، ۷ و ۹. مقاله حاضر سعی دارد تا المان‌های شخصی‌سازی را به سیستم‌های توصیه‌گری که پیشنهادات آنها مبتنی بر پالایش مشارکتی و شباهت کاربران است اضافه نماید، تا با افزودن میزان ارتباط میان اشخاص و علایق‌شان خدمات مناسب‌تری را فراهم سازد. لذا در این مقاله از شبکه باور عمیق (DBN)^۲ به منظور رویکردی با قابلیت پردازش و تفکیک ویژگی‌های متعدد و مرتبط با داده‌های زیاد استفاده شده است. در زمینه سیستم‌های توصیه‌گر گردشگری مطالعاتی انجام شده است که مواردی از آنها بررسی شده است.

گارسیا و همکاران (۲۰۱۱) با هدف ارائه پیشنهاداتی جهت فعالیت‌های گردشگری، برای تورهای گردشگری براساس ارجحیت‌های کاربران، خصوصیات دموگرافیک و همچنین نوع مکان‌هایی که در سفرهای پیشین بازدید کرده بودند یک سرویس مبتنی بر وب در شهر والنسیا^۳ اسپانیا طراحی نمودند / ۱۰. گوالاس و کنتریس (۲۰۱۱) یک برنامه کاربردی بافت آگاه تحت موبایل برای برنامه سفر روزانه شهروندان ارائه دادند

^۲ Deep Belief Network^۳ Valencia^۱ Collaborative Filtering

کرستانی (۲۰۱۸) یک سامانه توصیه‌گر مکان را برای گردشگری پیاده‌سازی کرده‌اند. در عملکرد این سیستم داده‌های محتوامحور، برچسب‌ها و شناسایی موقعیت مکان در نظر گرفته شده‌است. سیستم توصیه‌گر مذکور، مکان‌های توصیه شده را براساس شبکه‌های اجتماعی مکان مبنای فوراسکوئر^۳ و یلپ^۴ به گردشگران معرفی می‌کند. در این پژوهش، یک مدل احتمالی برای پیدا کردن نگاشت بین برچسب‌های ارسالی کاربران و کلمات کلیدی پیشنهاد شد. این نگاشت، آنها را قادر ساخت تا از جهت‌های مختلف برای رفع مشکل پراکندگی داده‌ها برای توصیه‌گری استفاده کنند. نتایج حاصل نقش داده‌های نزدیک‌تر به زمان تصمیم‌گیری را در به‌کارگیری روش‌های محتوامحور نشان می‌دهد / ۱۵. شیانگنان و همکاران (۲۰۱۸) برای توصیه جاذبه‌های گردشگری با روش‌های یادگیری ماشین از داده‌های موجود در وبسایت‌های ارتباط جمعی استفاده کردند. در این مطالعه یک سیستم توصیه مبتنی بر مدل کاربر پیشنهاد شده‌است که یک معماری دو مرحله‌ای برای تطبیق و رتبه‌بندی دارد. برای فرآیند تطبیق، یک مدل ماشین بردار به‌کار گرفته شده‌است. در نتیجه کاندیدها از طریق وارد نمودن تاریخ‌های سفر به کاربر ایجاد می‌شوند. نتایج نشان می‌دهد که این مدل با توجه به دقت و توانایی رتبه‌بندی بهتر از روش‌های مرسوم است. این مدل کمتر تحت تاثیر مشکل شروع سرد می‌باشد. باین‌حال، به دلیل محدودیت در کمیت و کیفیت مجموعه داده‌ها روش پیشنهادی می‌تواند عملکرد بهتری داشته‌باشد / ۱۶. کلاهکج و همکاران (۲۰۱۹) در سیستم توصیه‌گری خود از روشی به‌نام "پالایش گروهی مبتنی بر محتوای دموگرافیک" برای بررسی مشکلات ذکر شده استفاده کردند. در این پژوهش اطلاعات شخصی و جانبی در ترکیب با

سیستم مذکور برای حمایت از انتقال گردشگران پیاده یا در حال حرکت با استفاده از وسایل نقلیه عمومی و با در نظر گرفتن تاخیرهای موجود در ایستگاه‌های حمل و نقل توسعه داده شده‌است. در این سیستم، حالت‌های مختلف سفر بعنوان ارجحیت شخصی و بافت‌های اصلی در نظر گرفته شد / ۱۱. نیسانی‌سامانی و همکاران (۲۰۱۵) یک سیستم کاربرمحور گردشگری را جهت توصیه نقاط مورد علاقه براساس شرایط آب و هوایی پیاده‌سازی کرده‌اند. در این مطالعه از روش میانگین وزن دار ترتیبی (OWA)^۱ به عنوان تکنیک تصمیم‌گیری چندمعیاره (MCDM)^۲ استفاده شده است. نتایج تجربی در یک سناریوی راهیابی گردشگر با توجه به رضایت کاربران در ۱۰۰ تکرار الگوریتم با ۱۰۰ گردشگر با ترجیحات مختلف در شهر تهران ارزیابی شده‌است و نتایج حاصل اهمیت فرآیند وزندهی محلی را در ارائه پیشنهادات سیستم توصیه‌گر نشان می‌دهد / ۱۲. اسماعیلی و همکاران (۲۰۱۵) سیستم توصیه‌گر بافت آگاهی را جهت پیشنهاد مناسب‌ترین محل برای اقامت گردشگران شهر کرمان طراحی و پیاده‌سازی نمودند. سیستم مذکور با توجه به بافت و شرایط کاربر نظیر میزان هزینه مورد نظر و امکانات رفاهی بهترین مکان را از بین اقامتگاه‌های موجود در پایگاه داده به کاربر توصیه می‌کند / ۱۳. گودرزی و همکاران (۲۰۱۵) در مطالعه-ای الگوریتم توسعه مکان را از لحاظ تاثیر مکانی، زمانی و روابط اجتماعی توسعه داده‌اند. در توسعه این الگوریتم، از تابع توزیع مکانی برای مدل‌سازی تاثیر مکانی، از منحنی زمانی هر کاربر در هر دسته برای مدل‌سازی تاثیر زمانی و برای مدل‌سازی تاثیر روابط اجتماعی از سه معیار شباهت ارتباطات اجتماعی، فعالیت‌های ثبت موقعیت و شباهت زمانی در رفتار ثبت موقعیت استفاده شده است / ۱۴. علی‌نژادی و

^۳ Foursquare

^۴ Yelp

^۱ Ordered Weighted Averaging

^۲ Multi-Criteria Decision Making

بازخوردهای کاربر برای غلبه بر محدودیت‌های روش‌های فیلترینگ هم‌کاری و در برخورد با داده‌های چندبُعدی موردتوجه قرار گرفت. با وجود مزایای مذکور از لحاظ دقت و صحت نسبت به روش‌های دیگر، روش پیشنهادی دارای نقاط ضعف متعددی است. اول، توالی نواحی برای توصیه در مورد هدف سفر برنامه‌ریزی نشده است. دوم، محدودیت‌های زمانی و بودجه‌ای به دلیل نیاز به پرسش‌های صریح از کاربران برای هر توصیه و در دسترس نبودن این اطلاعات در مجموعه داده مورد استفاده در نظر گرفته نشده است / ۱۷. گونگلی و همکاران (۲۰۱۹) سیستم توصیه ترکیبی برای نقاط گردشگری را با استفاده از آمار نمونه‌برداری سلسله مراتبی و رتبه‌بندی شخصی چند وجهی پیشنهاد دادند. این تحقیق روی طراحی یک سیستم توصیه ترکیبی جدید برای نقاط گردشگری متمرکز شدند. از تصاویر نقاط گردشگری برای حل مساله تنگی داده‌ها در این روش استفاده شده است. نتایج تجربی ایشان نشان می‌دهد که سیستم توصیه ترکیبی پیشنهادی برای نقاط گردشگری کارآمد و مقاوم است. از آن مهم‌تر، سیستم توصیه ترکیبی پیشنهادی در توصیه یک گروه از نقاط گردشگری ثبات بیشتری دارد که ارزش عملی بالای آن را نشان می‌دهد. با این حال، سیستم توصیه ترکیبی هنوز هم می‌تواند بهبودیابد / ۱۸. عباسی مود و همکاران (۲۰۲۱) روشی را برای تشخیص اطلاعات معنایی مناطق گردشگری با استفاده از تحلیل احساسات^۱ افراد در مواجهه با مکان‌های گردشگری ارائه دادند. همچنین این سیستم از اطلاعات حیاتی متنی زمان، مکان و آب و هوا برای فیلتر کردن موارد نامناسب و افزایش کیفیت پیشنهادات در مورد وضعیت فعلی استفاده نمود. نتایج حاصل افزایش رضایت کاربران را نشان می‌دهد / ۱۹. چن و هندری (۲۰۱۹) از یک شبکه عصبی باور عمیق برای شخصی‌سازی یک

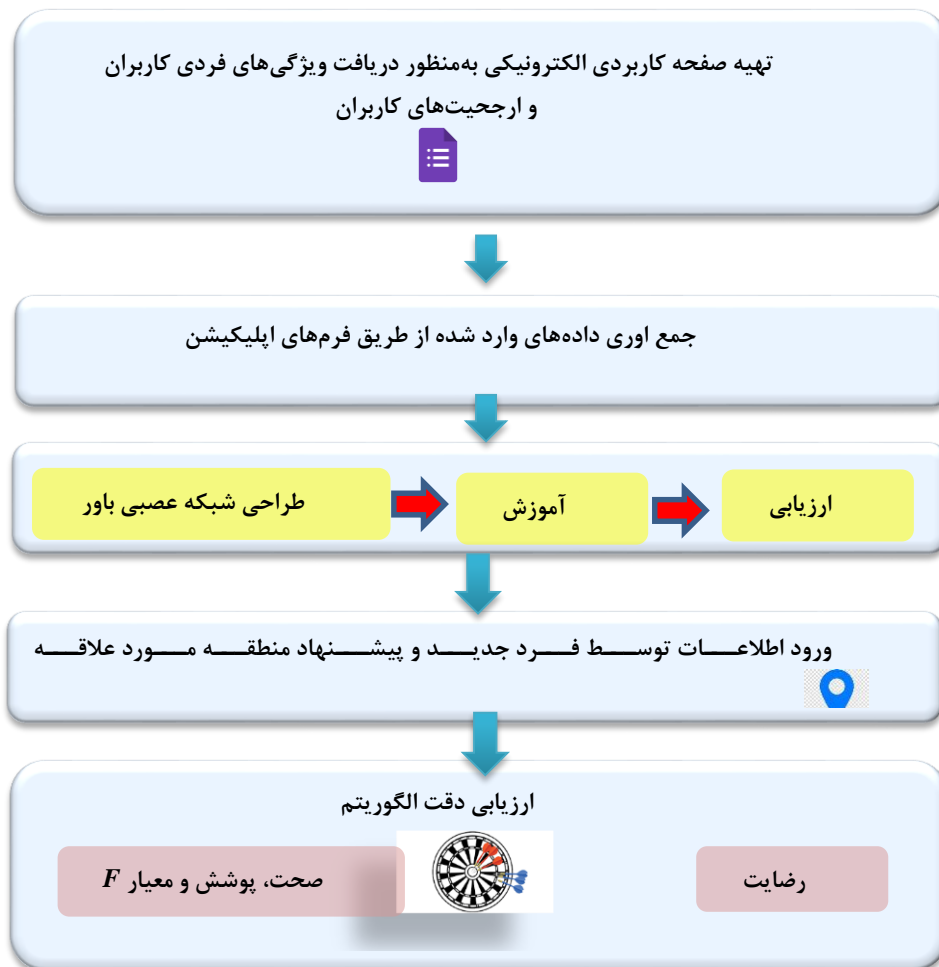
سیستم توصیه‌گر براساس کامنت‌های ارسالی از سوی کاربران به صورت محتوامحور استفاده نمودند. در این تحقیق رتبه‌بندی توصیه‌ها براساس تناسب با کاربر انجام شد. نتایج حاصل کارایی این شبکه را برای متناسب‌سازی و پیش‌بینی توصیه‌های مناسب نشان می‌دهد / ۲۰. قاسمی و همکاران (۲۰۱۶) از شبکه عصبی باور عمیق برای کلاسه‌بندی داده‌های نامتعادل استفاده نمودند و به این نتیجه دست یافتند که تغییر پارامترهای شبکه، نقش مهمی در ارائه خروجی‌ها با دقت بالا فراهم می‌سازد / ۲۱. ژانگ و لیو (۲۰۲۰) به توسعه شبکه عصبی برای پیش‌بینی الگوی ترافیکی شهر پرداختند و مقایسه نتایج حاصل با دیگر الگوریتم‌های خوشه‌بندی، کارایی روش توسعه داده‌شده را نشان داد / ۲۲.

بررسی تحقیقات گذشته نشان می‌دهد که اگرچه مباحث مربوط به مشابهت کاربران در یافتن مکان‌های موردعلاقه کاربران بسیار اهمیت دارد ولی به کارگیری روش‌های محتوامحور نیز می‌تواند در فرایند شخصی‌سازی سیستم‌های توصیه‌گر موثر باشد. از این رو مقاله حاضر با توجه به تعداد ویژگی‌های متعدد شخصی‌سازی، رویکرد جدیدی را با به کارگیری شبکه‌های باور عمیق و دقیق‌ترین مکان‌های مطابق با نیازهای کاربران ارائه می‌دهد. تفاوت اصلی این مقاله با مطالعات مطرح در این حوزه، شخصی‌سازی سیستم‌های توصیه‌گر گردشگری به صورت محتوامحور و با به کارگیری شبکه باور عمیق به عنوان یک رویکرد قابل توسعه است. ضمن اینکه با افزایش داده‌ها در طول زمان و آموزش مجدد شبکه دقت خروجی‌ها نیز افزایش می‌یابد.

۲- روش تحقیق

روش پیشنهادی این تحقیق دارای پنج مرحله اصلی است که در شکل (۱) نمایش داده شده است.

^۱ Sentiment Analysis



شکل ۱: دیگرام روش تحقیق پیشنهادی

شاخص های صحت^۱، پوشش^۲ و معیار F ^۳ و رضایت کاربران انجام خواهد شد.

۲-۱- طراحی شبکه باور عمیق

در این مقاله برای فرآیند شخصی سازی توصیه های گردشگری از شبکه باور عمیق استفاده شده است. علت به کارگیری این شبکه امکان تولید داده های مشابه با داده های واقعی برای آموزش شبکه است. از سوی دیگر برای تفکیک خصوصیات و ارجحیت های کاربران با

در مرحله اول صفحه الکترونیکی جمع آوری داده ها و ارجحیت های کاربران طراحی شده و در اختیار گردشگران قرار می گیرد. در مرحله دوم داده های مربوط به ویژگی های فردی کاربران و داده های مربوط به ارجحیت های کاربران و نقاط مورد علاقه آنها از کاربر اخذ می گردد. در مرحله سوم شبکه باور عمیق به منظور خوشه بندی کاربران مشابه و ارائه اطلاعات شخصی سازی شده تدوین و مورد آموزش واقع می شود. در مرحله چهارم اطلاعات افراد جدید وارد شبکه آموزش دیده شده و نقاط مورد علاقه کاربران جدید مشخص می شود. در مرحله نهایی ارزیابی دقت الگوریتم براساس

^۱ Precision

^۲ Recall

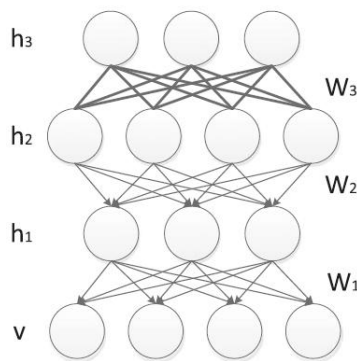
^۳ F-Score

آموزش *DBN* در دو مرحله انجام می‌شود:

(۱) پیش آموزش^۳

(۲) تنظیم دقیق^۴.

در مرحله اول استخراج ویژگی با انجام یک آموزش بدون نظارت در جهت پایین به بالا انجام می‌شود. درحالیکه در مرحله دوم یک الگوریتم نظارت از بالا به پایین برای تنظیم بیشتر پارامترهای شبکه اجرا می‌شود. علت استفاده از دو فاز آموزشی در مدل‌های عمیق، وجود تعداد بسیار زیادی لایه پنهان در شبکه است که در صورت آموزش یک مرحله‌ای موجب کم شدن گرادیان در لایه‌های اول (محو شدن گرادیان) و در نتیجه کاهش سرعت آموزش می‌شود که از مهم‌ترین مشکلات یادگیری عمیق است. به همین دلیل برای رفع این مشکل در ابتدا ماشین‌های بولترمن محدود تشکیل دهنده شبکه جداگانه آموزش می‌بینند و سپس آموزش بر کل شبکه اعمال می‌شود. فرمول بندی *RBM* در ادامه توضیح داده می‌شود.



شکل ۲: نمایی از لایه‌های شبکه باور عمیق

توجه به آموزش لایه به لایه شبکه امکان تفکیک با دقت بالاتری وجود خواهد داشت. شبکه باور عمیق نوعی مدل مبتنی بر گراف مولد عمیق یا نوعی شبکه عصبی عمیق است و شامل مجموعه‌ای از ماشین‌های بولترمن محدود (*RBM*) می‌باشد. برخلاف استفاده از وزن‌های تصادفی مزیت اصلی شبکه باور عمیق مقداردهی وزن اولیه یک شبکه عصبی عمیق به منظور دستیابی به یک مدل مطلوب است. شبکه‌های باور عمیق می‌توانند برای اجرای موفقیت‌آمیز آموزش لایه به لایه از الگوریتم پس انتشار خطا^۲ استفاده کنند. رویکرد متداول استفاده از مدل احتمالی مبتنی بر انرژی برای ایجاد توزیع مشترک بین داده‌های مشاهده شده "*v*" و متغیرهای پنهان "*h*" است (رابطه (۱)):

رابطه (۱)

$$p(v, h^1, \dots, h^m) = \left(\prod_{i=1}^{m-2} p(h^{i+1} | h^i) \right) \cdot p(h^{m-1} | h^m)$$

در رابطه (۱)، $h^0 = 1$ و $P(h^i | h^{i+1})$ یک توزیع شرطی برای واحدهای پنهان در *RBM* برای سطح *k* ام *DBN* است و $P(h^{m-1} | h^m)$ توزیع همزمان پنهان در بالاترین سطح *RBM* است. خروجی هر لایه به عنوان ورودی لایه بعدی اعمال می‌شود. شکل (۲) نمای کلی ساختار *DBN* ها را نشان می‌دهد (که در آن h_i ها لایه‌های پنهان و w_i وزن‌های متغیر هر لایه است) [۲۳].

مطابق شکل (۲)، یک *DBN*، برای هر دو لایه همسایه، یک *RBM* ساخته می‌شود. لایه قابل مشاهده هر *RBM* مربوط به لایه پنهان *RBM* قبلی است و دو لایه بالاتر غیر جهت‌دار هستند. اتصال جهت‌دار بین لایه‌های بالاتر و پایین‌تر از بالا به پایین است. آموزش چندین لایه *RBM* در *DBN* به طور متوالی انجام می‌شود. ابتدا *RBM*‌های پایینی و سپس بالاترین لایه‌ها آموزش داده می‌شوند. پس از تعیین مشخصات توسط *RBM* بالایی نتایج را می‌توان به لایه‌های پایین‌تر انتشار داد. روش

^۳ Pre-Training

^۴ Fine-Tuning

^۱ Restricted Boltzmann Machine

^۲ Backpropagation

معیارهای بررسی خروجی سیستم‌های توصیه‌گر هستند [۲۳]. از آنجایی که این دومعیار در تقابل با یکدیگر تغییر می‌کنند لذا معیار ارزیابی جدیدی تحت عنوان معیار F که بر مبنای ترکیب دو معیار مذکور کار می‌کند معرفی و استفاده شده است. برای محاسبه مقادیر این معیار ابتدا لازم است تعریفی از صحت و پوشش ارائه شود. معیار صحت، نسبت تعداد پیش‌بینی‌های صحیح انجام شده مطابق خواسته کاربر برای مناطق گردشگری، به تعداد کل مناطق پیشنهادی را ارزیابی می‌کند (رابطه (۵)):

$$Precision = \frac{TP}{TP + FP}$$

در رابطه (۵) TP^2 مثبت حقیقی یا همان نقاط پیشنهادی مورد علاقه کاربر است. FP^3 مثبت کاذب یا نقاط پیشنهادی نامناسب برای کاربر است. معیار پوشش بیان‌کننده نسبت تعداد پیش‌بینی‌های صحیح انجام شده برای مناطق گردشگری به تعداد مناطق درست پیش‌بینی شده به علاوه مناطقی که پیشنهاد نشده و درست هم نبوده است، می‌باشد (رابطه (۶)):

$$Recall = \frac{TP}{TP + FN}$$

FN^4 منفی کاذب یا همان نقاط پیشنهاد نشده و مناسب کاربر است. معیار F پارامترهای صحت و پوشش را با هم ترکیب می‌کند تا مشخص شود یک مدل توصیه‌گر تا چه حد عملکرد خوبی از خود نشان می‌دهد (رابطه (۷)):

۲-۲- ماشین بولتزمن محدود

ماشین بولتزمن محدود نوع خاصی از ماشین بولتزمن است که در آن هیچ ارتباط داخلی بین لایه‌های مرئی و پنهان وجود ندارد. احتمال پیکربندی مشترک $(1, h)$ برای RBM با توجه به رابطه (۲) محاسبه می‌شود [۲۳]:

$$Pr(1, h) = \frac{\exp(-Energy(1, h))}{Z}$$

در رابطه (۲) $Z = \sum_{i,j} \exp(-Energy(l_i, h_j))$ ، فاکتور نرمال‌سازی می‌باشد. احتمال یک واحد مرئی با جمع کردن تمام اجزای پنهان محاسبه می‌شود که با استفاده از رابطه (۳) قابل محاسبه است:

$$p(v) = \frac{1}{Z} \sum_h \exp(-Energy(1, h))$$

که $\phi +$ و $\phi -$ به ترتیب فازهای مثبت و منفی را نشان می‌دهند. یکی از الگوریتم‌های مناسب برای تخمین گرادین احتمال ورود، واگرایی انقباضی^۱ (CD) است. الگوریتم CD توانایی به‌روزرسانی پارامترهای آموزش، بایاس‌ها و وزن‌ها را در RBM دارد. در نهایت اگر احتمال بیشتر از حد آستانه باشد لایه پنهان روشن می‌شود. به‌طور کلی از احتمال pi در به‌روزرسانی واحدهای مرئی استفاده می‌شود (رابطه (۴)):

$$p(v_i) = b_i + \sum_j w_{i,j} h_j$$

بایاس‌ها و وزن‌ها با تقریب‌سازی گرادین به‌روز می‌شوند. به‌منظور پیاده‌سازی روش تحقیق پیشنهادی از دو نوع داده شامل اطلاعات کاربران برای تعیین مشابهت و داده‌های مربوط به عملیات شخصی‌سازی استفاده شد.

۲-۳- ارزیابی دقت الگوریتم

در ادبیات تحقیق دومعیار صحت و پوشش رایج ترین

^۲ True Positive

^۳ False Positive

^۴ False Negative

^۱ Contrastive Divergence

رابطه (۷)

$$F - Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

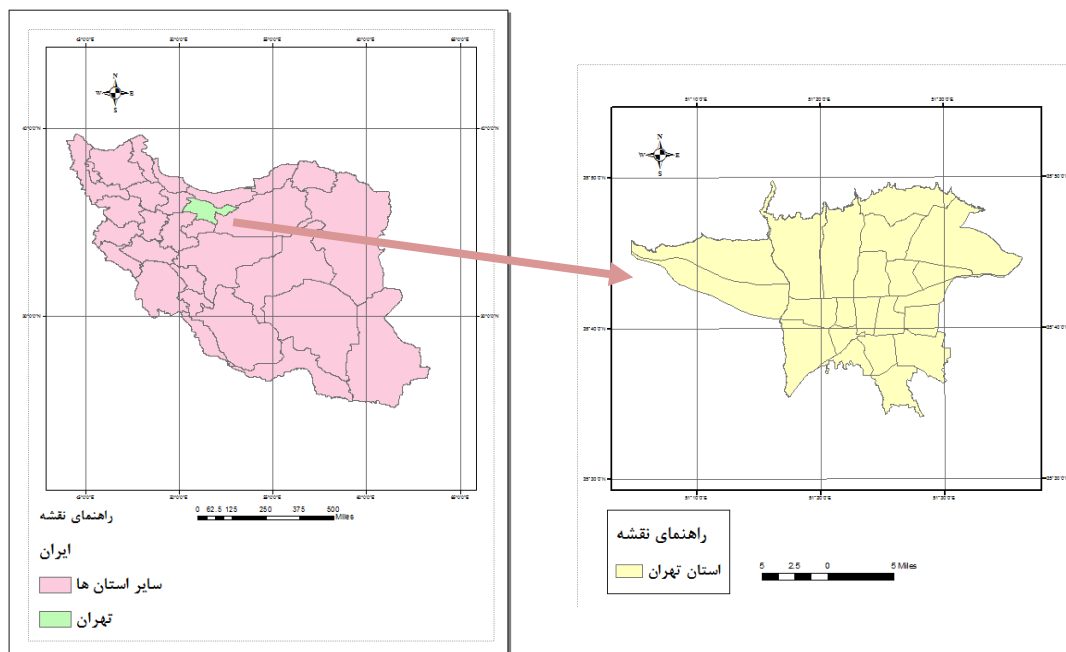
۳- پیاده‌سازی

در این بخش تمام مباحث مربوط به پیاده‌سازی الگوریتم مورد نظر ارائه می‌گردد:

۳-۱- منطقه مورد مطالعه

شهر تهران که پایتخت ایران نیز محسوب می‌شود بین ۳۴ تا ۳۶٫۵ درجه عرض شمالی و ۵۰ تا ۵۳ درجه طول شرقی و وسعت حدود ۱۲٫۹۸۱ کیلومتر مربع در

مرکز شهر تهران واقع است. طبق سرشماری سال ۱۳۹۵ استان تهران با بیش از ۱۳ میلیون نفر جمعیت، ۵٫۱۷ درصد جمعیت کل کشور را در خود جای داده‌است. در استان تهران ۶۶۳ درصد از جمعیت شهری در شهر تهران هستند. استان تهران از جنوب به استان قم، از شرق به استان سمنان، از شمال به استان مازندران، از غرب به استان البرز و از جنوب غربی به استان مرکزی محدود شده‌است [۲۵]. شکل (۳) منطقه مورد مطالعه را نشان می‌دهد.



شکل ۳: منطقه مورد مطالعه

۳-۲- معیارها و داده‌های مورد استفاده

مجموعه داده مورد استفاده در پژوهش حاضر با استفاده از توزیع پرسشنامه به تعداد ۴۰۰ عدد بین گردشگران جمع‌آوری شده‌است. در فرم مربوطه ویژگی‌های عوامل توصیه‌گری شامل اطلاعات فردی و همچنین ارجحیت‌های شخصی شامل ۱۴ المان جنسیت، سن، شغل، درآمد ماهیانه، وضعیت تاهل، محل سکونت، توانایی جسمانی، رشته ورزشی تخصصی، علاقه به گردشگری، وسیله مورد استفاده، توانایی پیاده‌روی، توانایی

کوهنوردی، نزدیکی به وسایل نقلیه عمومی و محدودیت زمانی (مدت اقامت) در نظر گرفته شد. در این تحقیق برای برآورد حجم نمونه از فرمول کوکران در سطح خطای ۵٪ استفاده شد که در نهایت با بکارگیری جدول مورگان براساس حجم جامعه توریستی ۴۰۰۰ نفر در منطقه مورد مطالعه حداقل حجم نمونه ۳۵۰ نفر محاسبه شد که به ۴۰۰ نمونه ارتقا یافت. فرمول کوکران یکی از روش‌های آماری است که معمولاً در ارتباط با مطالعه متغیرهای کیفی برای تعیین حجم

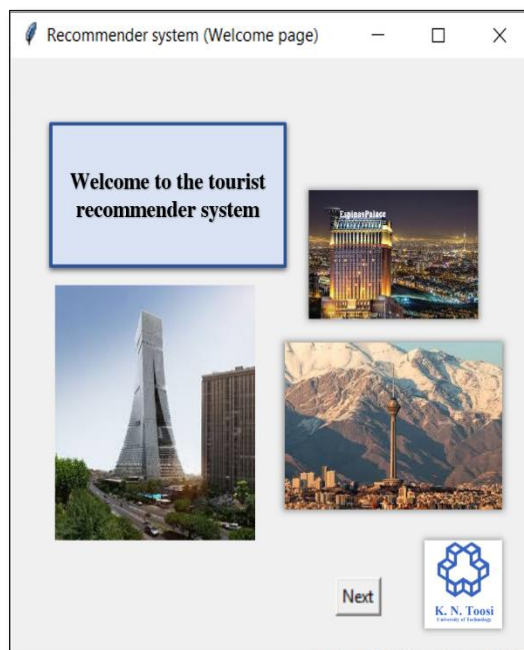
و Z^2 برابر $3/8416$ است. مقدار d نیز تفاضل نسبت واقعی صفت در جامعه با میزان تخمین برای وجود آن ویژگی در جامعه است. دقت نمونه گیری به این عامل بستگی دارد و اگر نمونه گیری دارای بیشترین دقت باشد از حداکثر مقدار d برابر 0.5 استفاده می شود [۲۶]. برای دریافت اطلاعات کاربران فرمهایی با زبان برنامه نویسی پایتون طراحی گردید. بر اساس ارتباط آن با پایگاه داده بصورت یک اپلیکیشن کاربردی در اختیار گردشگران منطقه موردنظر قرار گرفت. شکل (۴) نمونه هایی از فرم های الکترونیکی طراحی شده برای ورود اطلاعات مورد نیاز کاربران را نشان می دهد.

نمونه مورد استفاده قرار می گیرد. حجم نمونه در این روش مطابق رابطه (۸) محاسبه می شود:

رابطه (۸)

$$n = \frac{\frac{z^2 pq}{d^2}}{1 + \frac{1}{N} [\frac{z^2 pq}{d^2} - 1]}$$

در رابطه (۸)، N حجم جامعه، آماره p درصد توزیع ویژگی در جامعه یعنی نسبت افرادی است که دارای ویژگی مورد مطالعه هستند و آماره q نیز درصد افرادی است که فاقد ویژگی مورد مطالعه هستند. اگر میزان p و q مشخص نباشد از حداکثر مقدار آنها یعنی 0.5 استفاده می شود. در سطح خطای 5% مقدار z برابر 1.96



شکل ۴: نمونه هایی از فرم های الکترونیکی ورود اطلاعات مورد نیاز کاربران به منظور شخصی سازی

غیر واقعی تولید شده توسط شبکه مولد آموزش داده خواهد شد. در این مرحله مولد در حال توقف است و هیچ بازخوردی از آموزش دریافت نمی کند و تنها تفکیک کننده است که آموزش می بیند. به عبارت دیگر خطا در شبکه مولد پس انتشار نمی شود. در نتیجه

۳-۳- طراحی و آموزش شبکه باور عمیق

شروع کار با آموزش تفکیک کننده روی داده های واقعی برای چند بازه زمانی آغاز گردید. بدین منظور که تفکیک کننده به داده های واقعی ارزش بالاتری اختصاص دهد. سپس همان شبکه روی داده های

لایه ورودی، سه لایه مخفی و یک لایه خروجی) طراحی گردید و در مناسب‌ترین حالت همگرایی الگوریتم در ۲۰ بار تکرار اجرا شد. تعداد ۲۵ نورون در لایه‌های پنهان در نظر گرفته شد. به منظور مشخص نمودن چگونگی تعیین پارامترهای شبکه دقت، شبکه‌های طراحی شده برای تعداد نورون‌ها و تعداد تکرارهای مختلف در جدول (۱) آمده است. همانطور که دیده می‌شود سطر سوم بهترین پاسخ را نمایش می‌دهد.

مراحل قبل شبکه تفکیک کننده به طور قابل توجهی به نسبت شبکه مولد که تا کنون هیچ آموزشی ندیده است و همچنان نویز تولید می‌کند، در کار خود بهتر است. بنابراین تفکیک کننده را در حالت توقف قرار داده و شبکه مولد با استفاده از بازخورد تفکیک کننده، آموزش داده شد. این آموزش متناوب بین دو شبکه را تا زمانی که نتایج خوبی روی داده‌های تولید شده حاصل شود، ادامه یافت. براساس خطای جذر میانگین مربعات ($RMSE$) بدست آمده بهترین معماری شخصی سازی، یک شبکه باور عمیق با پنج لایه (یک

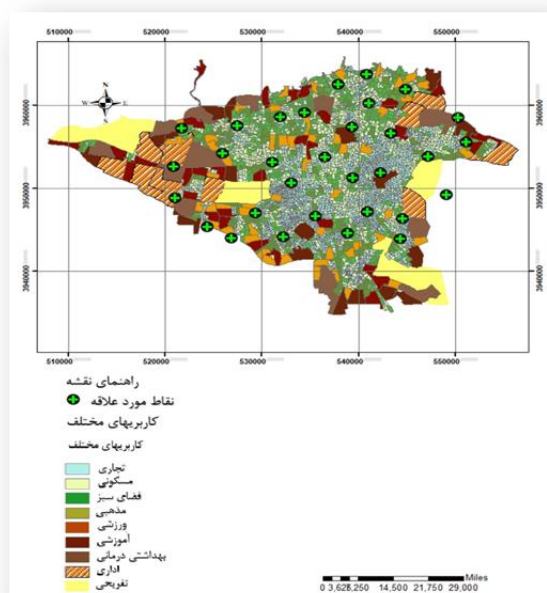
جدول ۱: تغییر پارامترهای شبکه و برآورد دقت شبکه

تعداد نورونها	تعداد تکرار	$RMSE$	R
۲۰	۳۲	۰/۰۰۲۵۴۷۸	۰/۹۸۵۴
۲۵	۲۰	۰/۰۰۰۳۲۵۴۷	۰/۹۹۸۷
۳۰	۱۸	۰/۰۰۴۵۸۷۱	۰/۹۸۸۴

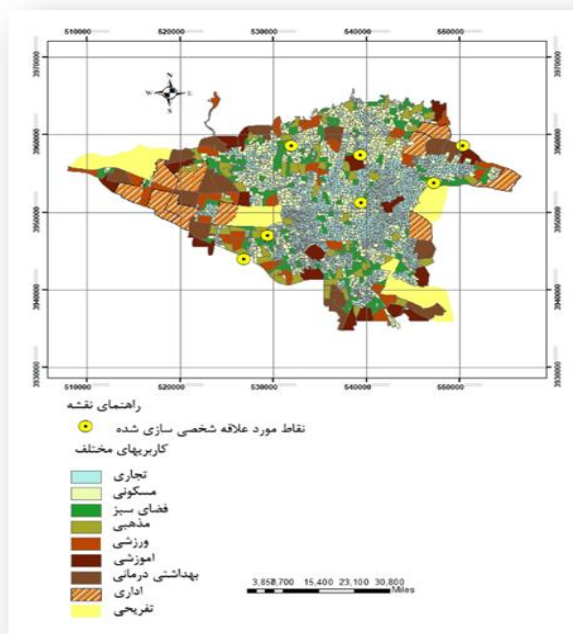
استفاده از الگوریتم شخصی سازی و با استفاده از آن برای دو شخص متفاوت نشان می‌دهد.

۴- بررسی نتایج

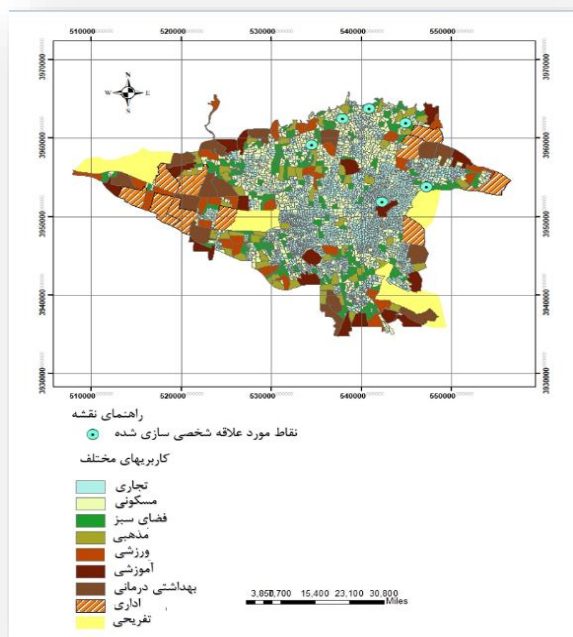
شکل (۵) نقاط مورد علاقه گردشگری را در حالت بدون



(الف)



(ب)



(ج)

شکل ۵: الف) کل نقاط مورد علاقه پیشنهادی، ب) نقاط مورد علاقه شخص اول، ج) نقاط مورد علاقه شخص دوم

۵- نتیجه گیری

امروزه با گسترش فناوری اطلاعات، امکان جستجو و دسترسی به اطلاعات مورد نیاز کاربران از طریق ارائه پیشنهادات در قالب سیستم های توصیه گر فراهم شده است. در این راستا نرم افزارهای گوناگونی در حوزه اطلاعات مکانی و غیر مکانی تولید شده است. با این وجود اغلب برنامه های ارائه شده به صورت توسعه یک مکان گردشگری می باشد و به مفاهیم شخصی سازی توجه زیادی نشده است. صنعت گردشگری سومین صنعت مولد جهان به شمار می رود و یکی از مهم ترین بخش های اقتصادی کشورها محسوب می شود. در این تحقیق به معرفی یک الگوریتم توصیه گر شخصی پرداخته شده است. الگوریتم ارائه شده توانایی توصیه مکان های گردشگری به صورت شخصی را دارد. همچنین در این مقاله اهمیت گردشگر و گردشگری نیز لحاظ شده است. سناریوی تحقیق به این صورت است که ابتدا گردشگران مختلف در سیستم ثبت نام می کنند، سپس اطلاعات فردی و ارجحیت های عمومی و عوامل خاص شخصی سازی خود را جهت بازدید از مراکز گردشگری بیان می کنند. در رویکرد پیشنهادی نیازی به تفکیک کاربران نیست بلکه با توجه به قدرت یادگیری شبکه های باور عمیق امکان تفکیک و شخصی سازی پیشنهادات کاربران وجود دارد. در این راستا داده های مربوط به ۴۰۰ مسافر گردشگری براساس ۱۴ المان ورودی و متمایز کننده دریافت شد. همچنین براساس شبکه آموزش دیده قابلیت پیش بینی مکان های گردشگری شخصی سازی شده برای ۳۰ کاربر جدید بررسی گردید. ارزیابی نتایج براساس سه شاخص صحت، پوشش و معیار F به ترتیب دقت میانگین ۷۱/۵۴٪، ۶۴/۱۶٪ و ۶۸/۴۳٪ را به همراه رضایت مندی بیش از ۷۹ درصدی کاربران نشان می دهد.

مقاله حاضر دارای ویژگی هایی است که آن را از مطالعات دیگر متمایز می سازد:

(۱) عموماً در فرآیندهای شخصی سازی از دو یا چند الگوریتم استفاده شده است که عموماً یکی از الگوریتم ها

به خوشه بندی و مشابهت یابی کاربران و الگوریتم دیگر به انجام فرآیند شخصی سازی می پردازد / ۱۷، ۱۶، ۱۵ و ۱۸ اما در روش پیشنهادی باتوجه به ویژگی آموزش لایه به لایه، امکان تشخیص و توصیه نقاط مورد علاقه فراهم است.

(۲) روش پیشنهادی باتوجه به داده مبنا بودن روش، امکان ارائه توصیه های مناسب را در صورت داشتن داده مناسب از منطقه دارد.

(۳) باتوجه به ویژگی های شبکه باور عمیق با افزایش حجم داده امکان ارائه توصیه های شخصی سازی شده زمان مند نیز وجود دارد که در تحقیقات آتی با فراهم نمودن داده زمانی با حجم بالا به آن پرداخته خواهد شد. این مدل روی داده هایی که خصوصیات و ارجحیت ها و نقاط مورد علاقه کاربران مشخص باشد قابل پیاده سازی است. محدودیت اصلی این روش نیاز به داده های زیاد برای پردازش با دقت بالا است که در تحقیقات بعدی چگونگی تامین داده از منابع متن باز و مردم گستر مورد توجه قرار خواهد گرفت.

روش پیشنهادی این تحقیق با همکاری ۳۰ کاربر آزمایش شده است. تمام کاربران قصد رفتن به چند جاذبه گردشگری را داشتند و از آنها خواسته شد تا به سوالات موجود پاسخ دهند. مقادیر محاسبه شده صحت و پوشش در جدول (۲) نمایش داده شده است.

به عنوان معیار ارزیابی سیستم های توصیه گر میانگین مقادیر پوشش، صحت و معیار F محاسبه و در جدول (۳) ارائه شده اند. هرچه معیار F به ۱۰۰ درصد نزدیکتر باشد به معنی صحت بیشتر خروجی و نزدیک بودن آن به خواست کاربر است. غالباً مقادیر قابل قبول برای معیار F در تحقیقات پیشین در زمینه سیستم های توصیه گر در بازه ۶۵٪ تا ۸۵٪ است.

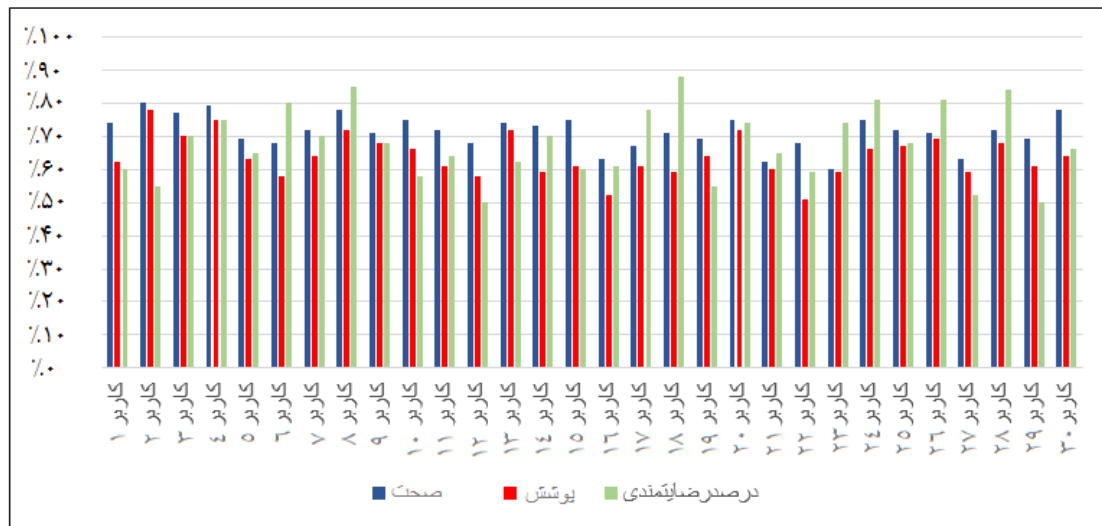
باتوجه به شکل (۶) نمودار مقایسه درصد مقادیر رضایت مندی، پوشش و صحت آمده است. همانطور که مشهود است تفاوت بسیار زیادی باتوجه به نظر سنجی کاربران بین مقادیر دیده نمی شود.

جدول ۲: مقادیر محاسبه شده صحت و پوشش

کاربر	صحت	پوشش
۱	٪۷۲	٪۶۳
۲	٪۸۲	٪۷۷
۳	٪۷۸	٪۶۹
۴	٪۷۸	۷۴
۵	٪۶۸	٪۶۲
۶	٪۶۹	٪۵۹
۷	٪۷۱	٪۶۳
۸	٪۷۹	٪۷۳
۹	٪۷۲	٪۶۹
۱۰	٪۷۴	٪۶۵
۱۱	٪۷۱	٪۶۰
۱۲	٪۶۹	٪۵۹
۱۳	٪۷۳	٪۷۱
۱۴	٪۷۴	٪۶۰
۱۵	٪۷۴	٪۶۲
۱۶	٪۶۴	٪۵۱
۱۷	٪۶۸	٪۶۲
۱۸	٪۷۲	٪۶۰
۱۹	٪۷۰	٪۶۳
۲۰	٪۷۴	٪۷۳
۲۱	٪۶۳	٪۶۱
۲۲	٪۶۸	٪۵۰
۲۳	٪۶۱	٪۵۸
۲۴	٪۷۴	٪۶۷
۲۵	٪۷۳	٪۶۸
۲۶	٪۷۰	٪۶۸
۲۷	٪۶۵	٪۵۸
۲۸	٪۷۰	٪۶۹
۲۹	٪۶۸	٪۶۰
۳۰	٪۷۹	٪۶۵

جدول ۳: میانگین پارامترهای صحت، پوشش و معیار F

میانگین	کمیت سنج
۷۱٫۵۴٪	صحت
۶۴٫۱۶٪	پوشش
۶۸٫۴۳٪	معیار F



شکل ۶: نمودار مقایسه درصد مقادیر رضایت مندی، صحت و پوشش

مراجع

- [1] N. Moshtaghi, H. Vahdat nejad, and M. Ghasemi Gol, "Provide context-aware tour guide system for group tourists," *Journal of Tourism Planning and Development*, 6, no. 23, pp. 59-76, 1396 (Persian).
- [2] A. Heidari Nik, and M. Talei, "Development of a Web GIS-Based recommendation System for Tourism," *Journal of Space Planning and Planning (Lecturer of Humanities. Fall 2012)*, 16. no 3 (consecutive 75); pp. 141-159, 1391 (Persian).
- [3] N. Neysani Samany, M.R. Delavar, N. Chrisman, and M.R. Malek, "An ontology for Spatial Relevant Objects in a Location-aware System: Case study: A tourist guide system", *World Academy of Science, Engineering and Technology*, vol. 63, pp. 878-884, 2012.
- [4] G.M. Harshvardhan, M.K. Gourisari, S.S. Rautaray, and M. Pandey, "UBMTR: Unsupervised Boltzmann machine-based time-aware recommendation system", *Journal of King Saud University - Computer and Information Sciences*, in press, 2021.
- [5] J. Bobadilla, F. Ortega, A. Hernando, and J. Alcalá, "Improving collaborative filtering recommender system results and performance using genetic algorithms", *Knowledge-Based Systems*, vol. 24, pp. 1310-1316, 2011.
- [6] H.S. Chiang, and H. Tien-Chi, "User-adapted travel planning system for personalized schedule recommendation", *Information Fusion*, vol. 21, pp. 3-17, 2015.
- [7] X. Cai, Z. Hu, P. Zhao, W. Zhang, and J. Chen, "A hybrid recommendation system

- with many-objective evolutionary algorithm", *Expert Systems with Applications*, vol.159, 2020.
- [8] K. Kabassi, "Personalizing recommendations for tourists", *Telematics and Informatics*, vol.27, pp. 51-66, 2010.
- [9] D. Almazro, G. Shahatah, L. Albdulkarim, M. Kherees, R. Martinez, and W. Nzoukou, "A Survey Paper on Recommender Systems", *ArXiv*, 2010.
- [10] I. García, L. Sebastia, and E. Onaindia, "On the design of individual and group recommender systems for tourism", *Expert Systems with Applications*, vol.38, pp. 7683-7692, 2011.
- [11] D. Gavalas, and M. Kenteris, "A web-based pervasive recommendation system for mobile tourist guides", *Personal and Ubiquitous Computing*, vol.15, pp. 759–770, 2011.
- [12] Z. Neisani, M. Torabi Azad, and N. Neysani Samani, "Design and Implementation of a User-Centered Tourism Guide System based on Weather Conditions Preferences", *2nd International Conference on Education, Management and Information Technology (ICEMIT)*, in Jinan, China. 2015.
- [13] Sh. Ismaili, M. R. Malek, f. Nasery, "Designing an informed context recommendation system to select the desired tourist accommodation," *The First International Conference on Human, Architecture, Civil Engineering and the City, Tabriz*, 2015 (Persian).
- [14] p. Goodarzi, A. A. Alesheikh, and S. Honarparvar, "Development of probabilistic place recommender algorithm based on location classification in location-based social networks based on temporal impact, spatial impact, and impact of social relations", *Journal of Surveying Science and Technology*, vol.7 . no: 2, pp. 215 – 230 , 1396(Persian).
- [15] M. Aliannejadi, and F. Crestani, "Personalized Context-Aware Point of Interest Recommendation", *ACM Transactions on Information Systems*, vol.36. no 4, pp.1-28, 2018.
- [16] He. Xiangnan, He. Zhankui, Du. Xiaoyu, Ch. Tat-Seng, "Adversarial Personalized Ranking for Recommendation", *Conference: The 41st International ACM SIGIR Conferenc*, pp. 355-364, 2018, DOI: 10.1145/3209978.3209981.
- [17] M. Kolahkaj, A. Harounabadi, A. Nikravanshalmani, and R. Chinipardaz, "DBCACF: A Multidimensional Method for Tourist Recommendation Based on Users", *Demographic, Context and Feedback*, vol.24, pp. 209-219 , 2019(persian).
- [18] G. Gungli, J. Hua, T. Yuan, J. Wu, Z. Jiang, H. Zhang, and T. Li, "Novel Recommendation System for Tourist Spots Based on Hierarchical Sampling Statistics and SVD++", *Mathematical Problems in Engineering*, pp. 1-15, 2019.
- [19] Z. Abbasi-Moud, H. Vahdat-Nejad, and J. Sadri , "Tourism recommendation system based on semantic clustering and sentiment analysis", *Expert Systems with Applications*, vol167, 2021.
- [20] R. Chen and H. Hendry , " User Rating Classification via Deep Belief Network Learning and Sentiment Analysis," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 3, pp. 1-12, 2019.
- [21] Y. Zhang and F. Liu, " An Improved Deep Belief Network Prediction Model Based on Knowledge Transfer", *Future Internet*, vol. 9, 2020; <https://doi.org/10.3390/fi12110188>.
- [22] Chen, C., Wang, H., Yuan, F, "Bus travel time prediction based on deep belief network with back-propagation", *Neural Comput & Applic*, vol. 32, pp.10435–10449, 2020. <https://doi.org/10.1007/s00521-019-04579-x>.
- [23] F. Ghasemi, A. Fassihi, H. Pérez-Sánchez, and A. Dehnavi, "The role of different

sampling methods in improving biological activity prediction using deep belief network”, Journal of computational chemistry , vol.38, pp. 195-203, 2016.

[24] A. Gunawardana, and G. Shani, “A Survey of Accuracy Evaluation Metrics of recommendation Tasks”, *Journal of Machine Learning Research*, vol.10, 2009.

[25] Department Of Environment Islamic Republic of Iran [Online]. Available: . <https://tehran.doe.ir/portal/home> , (1400-06-19).

[26] Cochrane, L., and Zeleny, M. (editors), “Multiple Criteria Decision-making”, (Columbia: University of South Carolina Press], 1973.



Personalization of a tourism recommender system based on users similarity and the use of deep belief network

Zeinab Neisani Samani ¹, Ali Asghar Alesheikh ^{2*}, Abolghasem Sadeghi-Niaraki ³, Mahdi Nazari Ashani ⁴

1- PhD student in Department of Geodesy and Geomatics Engineering, K.N. Toosi University of Technology

2- Professor in Department of Geodesy and Geomatics Engineering, K.N. Toosi University of Technology

3- Assistant professor in Department of Geodesy and Geomatics Engineering, K.N. Toosi University of Technology

4- Ms.c student of GIS in Department of Geodesy and Geomatics Engineering, K.N. Toosi University of Technology

Abstract

Spatial recommendation systems allow users to provide useful information by reducing duplicate and irrelevant information on the web widely. Recommendation systems are widely used in various fields, including tourism. Tourism recommendation systems can be used as tools by the tourist. A tourist can visit the tourist attractions of the region in a short time and with the least facilities, cost, and knowledge. Recommendation systems generally offer the necessary suggestions to different users based on participatory refinement and the similarity between the groups of the users. However, many services do not match the personal characteristics of the individual, and this reduces the effectiveness of such systems. The purpose of this study is to develop a recommendation algorithm based on the similarities between the users and personalization concepts. The innovation of this research is the use of a deep belief neural network to personalize the suggestions that can be offered to the tourists. The research scenario is as follows: first different tourists register in the system; then they express their personal information and general preferences and specific personalization factors for visiting the tourist centres. In the proposed approach, there is no need to separate the users; rather, due to the learning power of deep neural networks, it is possible to differentiate and personalize the user suggestions. In this regard, the related data to 400 tourists were received based on 14 input and distinguishing elements. Furthermore, based on the trained network, the predictability of personalized tourist places for 30 new users was examined. The results were evaluated based on these three indicators: Precision, Recall, and F-Score, as well as the user satisfaction. The results showed the high accuracy as well as the satisfaction of more than 79% of the users.

Key words: Recommender System, Tourism, Personalization, Deep Belief Network.