University of Qom

Engineering Management and Soft Computing

JEMSC.QOM
Online ISSN: 2538-2675

Developing a recommender system for the health tourism industry using data mining methods

Reza Molae Fard

Master of Computer-Software, Dezful Branch, Islamic Azad University, Dezful, Iran. Email: rezamolae4@gmail.com

Article Info

Article type:
Research Article

Article history:
Received 2022 August 27
Received in revised form 2022 October 2
Accepted 2022 October 6
Published online 2023 March 16

Keywords:
Recommender system, health tourism, data mining, participatory filtering.

ABSTRACT

Considering the depletion of oil resources in the coming years, the health tourism industry can be used as a suitable alternative to the country's source of income. To this reason, today, the health tourism industry is of great importance in developing countries such as Iran. However, this requires a lot of work, including preparation in terms of hardware as well as identifying places as the country's health tourism destinations accurately for foreign tourists who intend to visit our country. One way to correctly locate health tourism sites is to use recommender systems. Recommendation or recommender systems are those able to provide appropriate suggestions to users using limited information about their interests. In this research, a new method is presented to improve the recommendation systems in health tourism which can make accurate predictions using both participatory filtering and the points that previous tourists have given to places and health professionals in our country. According to research, data clustering using DBSCAN algorithm achieved a performance score of 99%, which is the highest among the existing algorithms. Moreover, the SVM method obtained a score of 95% in the accuracy section and a 99% score in the recall section, which proves the high accuracy of predicting the results. The proposed method, in general, can correctly locate up to 80% of the places needed by the tourist and suggest the appropriate place to a large extent.

Cite this article: Molae Fard, R. (2022). Developing a recommender system for the health tourism industry using data mining methods. *Enginiiring Management and Soft Computing, 8* (2). 1-18. **DOI**: https://doi.org/10.22091/jemsc.2020.5564.1134

© The Author(s)
DOI: https://doi.org/10.22091/jemsc.2020.5564.1134

Publisher: University of Qom

JEMSC.QOM
شاپا الکترونیکی: ۲۶۷۵-۲۵۳۸

مدیریت مهندسی و رایانش نرم

دانشگاه قم

ارائه یک سیستم توصیه‌گر برای صنعت گردشگری سلامت با استفاده از روش‌های داده‌کاوی

رضا مولایی‌فرد

کارشناسی ارشد کامپیوتر- نرم افزار، واحد دزفول، دانشگاه آزاد اسلامی، دزفول، ایران. رایانامه: rezamolae4@gmail.com

اطلاعات مقاله

نوع مقاله: مقاله پژوهشی

تاریخ دریافت: ۱۴۰۱/۰۶/۰۵

تاریخ بازنگری: ۱۴۰۱/۰۷/۱۰

تاریخ پذیرش: ۱۴۰۱/۰۷/۱۴

تاریخ انتشار: ۱۴۰۱/۱۲/۲۵

کلیدواژه‌ها:
سیستم توصیه‌گر،
گردشگری سلامت،
داده‌کاوی،
فیلترینگ مشارکتی.

چکیده

امروزه صنعت گردشگری سلامت در کشورهای درحال‌توسعه مانند ایران از اهمیت بالایی برخوردار است، زیرا با توجه به اتمام منابع نفتی تا سال‌های آینده، می‌توان از صنعت گردشگری سلامت برای جایگزینی مناسبی برای منبع درآمد کشور استفاده کرد، ولی این کار مستلزم انجام کارهای فراوانی ازجمله آمادگی ازلحاظ سخت‌افزاری و همچنین شناسایی مکان‌های گردشگری سلامت کشور به‌صورت دقیق برای گردشگران خارجی که قصد بازدید از کشورمان رادارند است. یکی از راه‌های شناسایی صحیح اماکن گردشگری سلامت، استفاده از سیستم‌های توصیه‌گر است. سیستم‌های توصیه‌گر یا پیشنهاددهنده سیستم‌هایی هستند که می‌توانند با گرفتن اطلاعات محدودی از علایق کاربر، پیشنهادات مناسبی را به وی ارائه دهند. در این تحقیق به ارائه روش جدیدی به‌منظور بهبود سیستم‌های توصیه‌گر در زمینه گردشگری سلامت پرداخته می‌شود که با استفاده از فیلترینگ مشارکتی و با استفاده از امتیازاتی که گردشگران قبلی، به مکان‌ها و متخصصین حوزه سلامت در کشورمان، داده‌اند می‌تواند پیش‌بینی‌های دقیقی را جهت استفاده گردشگران ارائه دهد. طبق تحقیقات صورت گرفته خوشه‌بندی داده‌ها با استفاده از الگوریتم DBSCAN، امتیاز کارایی ۹۹٪ را به دست آورد که بالاترین امتیاز کارایی در بین الگوریتم‌های موجود است، همچنین روش SVM در بخش دقت، امتیاز ۹۵٪ و در بخش فراخوانی، امتیاز ۹۹٪ را به دست آورد که نشان از دقت بالای پیش‌بینی نتایج را دارد و روش پیشنهادی به‌صورت کلی تا ۸۰٪ می‌تواند مکان‌های موردنیاز گردشگر را به‌درستی تشخیص داده و مکان مناسب را تا حدود زیادی به‌درستی پیشنهاد دهد.

استناد: مولایی‌فرد، رضا (۱۴۰۱). «ارائه یک سیستم توصیه‌گر برای صنعت گردشگری سلامت با استفاده از روش‌های داده‌کاوی». *مدیریت مهندسی و رایانش نرم، دوره ۸* (۲). صص: ۱۸-۱. https://doi.org/10.22091/jemsc.2020.5564.1134

ناشر: دانشگاه قم

© نویسندگان.

۱) مقدمه

صنعت گردشگری پزشکی ایران یکی از کم‌ریسک‌ترین و پرسودترین سرمایه‌گذاری‌ها محسوب می‌شود. گسترش صنعت گردشگری علاوه بر نقش و تأثیر آن در ابراز هویت ملی، موجب ارتقای ابعاد وسیع اقتصادی ازجمله ایجاد فرصت‌های شغلی، درآمدزایی، کاهش فقر و گسترش عدالت اجتماعی و رفاه در جامعه می‌شود. در میان حوزه‌های مختلف گردشگری سلامت به دلیل قابلیت و مزیت‌های رقابتی از توجه زیادی برخوردار شده است. گردشگری پس از صنعت نفت و خودروسازی، سومین صنعت بزرگ جهان به شمار می‌رود. این صنعت تأثیر بسزایی در درآمدهای ارزی کشورها دارد و به‌طور حتم در آینده ایی نزدیک به صنعت اول جهان تبدیل خواهد شد. با گذر زمان و پیشرفت‌های این صنعت، گردشگری از حالت عام خارج‌شده و به شاخه‌های تخصصی نظیر گردشگری فرهنگی، گردشگری ورزشی، گردشگری ماجراجویانه، گردشگری سلامت و... تقسیم‌شده است. آنچه مبنای این تقسیم‌بندی قرار می‌گیرد، نیت یا قصد اصلی گردشگر از گردشگری است. بنابراین، افرادی که به‌منظور برخورداری از خدمات سلامت به کشور دیگری سفر می‌کنند، در زمره گردشگران سلامت قرار دارند. برای توریسم درمانی که گاهی از آن به‌عنوان توریسم سلامت یا گردشگری سلامت نامبرده می‌شود، تعاریف گوناگونی بیان‌شده است. سازمان گردشگری جهانی[1] به‌طور خاص، گردشگری سلامت را چنین تعریف می‌کند؛ استفاده از خدماتی که به بهبود یا افزایش سلامتی و افزایش روحیه افراد (با استفاده از آب‌های معدنی، آب‌وهوا یا مداخلات پزشکی) منجر می‌شود و در مکانی خارج از محدوده سکونت فرد که بیش از ۲۴ ساعت است، به طول می‌انجامد. مقوله‌ی گردشگری سلامت فراتر از گردشگر درمانی است، آب‌های گرم و معدنی و لجن درمانی همراه با امکانات طبیعی نیز شامل مقوله‌ی گردشگری سلامت است. شعار جهانی توریسم سلامت، امکانات و خدمات در حد کشورهای جهان اول یا پیشرفته و قیمت‌ها و هزینه‌ها در حد کشورهای درحال‌توسعه و جهان سوم است. گردشگری سلامت از زمان یونان و روم باستان وجود داشته و سپس به بسیاری از کشورهای اروپایی و نقاط دیگر دنیا توسعه‌یافته است. از زمان باستان، افراد زیادی به‌منظور فراغت روحی و درمان به مجاورت رودخانه و آب‌های معدنی می‌رفتند. آب‌های گرم از قدیم‌الایام با آرامش تجدیدقوای جسمی و روحی انسان همراه بوده است. استفاده‌کنندگان ازاین‌گونه آب‌ها، قرن‌هاست که از آب‌های گرم استفاده می‌کنند و از آب‌های چشمه‌های معدنی می‌نوشند. در اروپا چشمه‌های آب‌معدنی و کلینیک‌های آب‌درمانی در اواخر قرن هجدهم موردتوجه جدی قرار گرفتند و علاوه بر این‌که به تجهیزات متناسب مجهز شدند فضاهای اطراف آن‌ها نیز برای بهره‌مندی از سکون و آرامش و چشم‌اندازهای طبیعی سامان یافتند. در این قبیل کلینیک‌ها گروه‌های مشاوره پزشکی برای تجویز مدت و نحوه استفاده از هر نوع آب برای درمان بیماری‌ها مشغول هستند. در آمریکا نیز مردم به چشمه‌های آب‌معدنی، اقامتگاه‌های مجاور دریا سفر می‌کردند. کشور ایران نیز با توجه به مکان‌های متعدد برای گردشگری سلامت و همچنین باوجود پزشکان متخصص و فوق متخصص و همچنین با توجه به هزینه کمتر نسبت به کشورهای اروپایی و آمریکایی می‌تواند مقصد جذابی برای گردشگران سایر کشورها باشد. صنعت گردشگری سلامت در کشور درحال‌توسعه‌ایی مانند ایران از اهمیت زیادی برخوردار است، زیرا علاوه بر آشنایی مردم سایر نقاط جهان بافرهنگ کشور ایران و مشاهده پیشرفت کشور ازلحاظ پزشکی و هم ازلحاظ اقتصادی نیز می‌تواند ارز بسیار زیادی را روانه کشور کند. با

[1] World trade organization

توجه به اتمام منابع نفتی تا سال‌های آینده، می‌توان از صنعت گردشگری سلامت برای جایگزینی مناسبی برای منبع درآمد کشور استفاده کرد، ولی این کار مستلزم انجام کارهای فراوانی ازجمله آمادگی ازلحاظ سخت‌افزاری و همچنین شناسایی مکان‌های گردشگری سلامت کشور به‌صورت دقیق برای گردشگران خارجی که قصد بازدید از کشورمان را دارند است. یکی از راه‌های شناسایی مناسب، استفاده از سیستم‌های توصیه‌گر است. سیستم‌های توصیه‌گر یا پیشنهاددهنده سیستم‌هایی هستند که می‌توانند با گرفتن اطلاعات محدودی از علایق کاربر، پیشنهادات مناسبی را به وی ارائه دهند. در این تحقیق یک سیستم توصیه‌گر گردشگری سلامت با استفاده از فیلترینگ مشارکتی و با استفاده از امتیازاتی که سایر گردشگرانی که قبلاً به کشور سفرکرده‌اند به مکان‌های سلامت و همچنین متخصصین پزشکی کشور داده‌اند، می‌تواند سالانه درصد زیادی از گردشگران خارجی را روانه کشور کند تا هم سایر مردم جهان بافرهنگ اصیل ایرانی و پیشرفت کشور در زمینه پزشکی آشنا شوند و هم منابع بسیار زیاد مالی را روانه کشور کند.

۲) پیشینه تحقیق

نیلاشی و همکاران در مقاله خود در سال ۲۰۱۷ به بررسی سیستم‌های توصیه‌گر در زمینه گردشگری سلامت با استفاده از فیلترنگ مشارکتی پرداختند. این محققان در مقاله خود به ارائه روش پیشنهادی جدیدی مبتنی بر CF چندمعیاری را پیشنهاد کردند تا دقت پیش‌بینی سیستم‌های توصیه‌گر در حوزه گردشگری را با استفاده از روش‌های خوشه‌بندی، کاهش ابعاد و روش‌های پیش‌بینی را افزایش دهند. آن‌ها از سیستم‌های استنتاج نوری- فازی متحرک و رگرسیون بردار پشتیبانی به‌عنوان تکنیک‌های پیش‌بینی، تجزیه‌وتحلیل مؤلفه اصلی یعنی PCA به‌عنوان یک روش کاهش ابعاد و نقشه‌های خودسازمانده و حداکثر سازی امید با EM و تکنیک‌های خوشه‌بندی استفاده کردند.

نگویان و ریکی در مقاله خود در سال ۲۰۱۸ به ارائه روشی به‌منظور توصیه در حوزه گردشگری پرداختند. این محققان با استفاده از رویکرد متفاوتی سعی در پیاده‌سازی روشی با استفاده از بحث‌های سایر گردشگران داشتند. در سیستم پیشنهادی وقتی اعضای گروه در مورد گزینه‌ها در حال بحث هستند، ترجیحات کاربران ثبت و از آن‌ها برای پیشنهاد به سایر کاربران استفاده می‌شود. سیستم پیشنهادی آن‌ها بر تعاملات کاربران نظارت کرده و مسیرها و توصیه‌های مناسبی را پیشنهاد خواهد داد. سیستم موردنظر بر روی یک گوشی هوشمند پیاده‌سازی و اجرا می‌شود و به‌عنوان یک کمک دهنده عمل می‌کند تا اعضای گروه بتوانند به توافق رسیده و یک تصمیم نهایی را بگیرند.

ژو و همکاران در مقاله خود در سال ۲۰۱۵ به ارائه روشی به‌منظور توصیه اماکن به گردشگران پرداختند. این محققان با در نظر گرفتن مدل‌سازی موضوع برای استخراج ترجیحات کاربر روشی را برای پیشنهاد مسیر گردشگری تفریحی و گردشگری سلامت بر اساس تراکم سوابق سفر سایر گردشگران به سایر شهرها ارائه دادند. این روش برای توصیه، از روش پالایش مشارکتی سیستم‌های توصیه‌گر بر اساس مشارکت جامعه در عکس‌هایی با برچسب جغرافیایی و محاسبه امتیاز کاربران استفاده کردند.

۳) گردشگری سلامت و انواع آن

ترکیبی از عوامل متعددی موجب رشد صعودی گرایش مردم به مسافرت‌های پزشکی شده است که ازجمله هزینه بالای

زندگی در کشورهای صنعتی، آسان شدن مسافرت‌های بین‌المللی و اصلاح و بهبود سطح تکنولوژی و استانداردهای پزشکی در بسیاری از کشورهای جهان، قابل‌ذکر هستند. یکی از دلایل کشش به‌سوی مسافرت پزشکی، راحتی آن در مقایسه با کشورهای دیگر است. در برخی کشورها که سیستم خدمات درمانی عمومی متداول است، معمولاً زمان زیادی برای پاسخگویی به نیاز شهروندان صرف می‌شود. بیماران ناگزیرند مدت طولانی در انتظار رسیدگی به وضعیتشان باشند. جوینده‌های درمان و سلامتی از هرکجا از دنیا می‌توانند به یک کشور واردشده و تحت خدمات و مراقبت‌های پزشکی در زمینه‌های چون سرطان، بیماری‌های مغز و اعصاب، جراحی‌های پیوند اعضاء، عمل‌های زیبایی و غیره قرار گیرند.

۴) گردشگری پزشکی

درواقع گردشگری پزشکی رایج‌ترین و حساس‌ترین زیرشاخه گردشگری سلامت است. این سفر عموماً برای درمان بیماری یا انجام عمل جراحی و یا چک‌آپ سلامت گردشگر در کلینیک ها و بیمارستان‌های کشوری با امکانات پزشکی بالا و هزینه‌های درمانی مناسب انجام می‌گیرد. گردشگری پزشکی یا گردشگری درمانی حساس‌ترین نوع گردشگری در بین تمام انواع گردشگری‌هاست، زیرا به‌طور مستقیم با جان و سلامت گردشگر مرتبط است. گرچه بسیاری از کشورهای توریستی دنیا امروزه برای کسب درآمد، خدمات گردشگری ارائه می‌دهند، اما مطمئناً توانمندی یک کشور در علم پزشکی و خدمات درمانی برخلاف گردشگری تفریحی و ورزشی و... چیزی نیست که صرفاً با ساخت‌وساز و صرف پول به دست می‌آید. وجود پزشکان متبحر و باتجربه، سابقه علم پزشکی، وجود زیرساخت‌های درمانی و نظارتی، وجود بیمارستان‌ها و کلینیک‌های استاندارد، تجهیزات پزشکی کامل و به‌روز، همچنین قوانین و نظارت‌های پزشکی جزء زیرساخت‌های مهم درمانی یک کشور به شمار می‌روند تا آن کشور بتواند در مراحل بعدی با ادغام خدمات اقامتی و گردشگری اقدام به فعالیت در زمینه گردشگری سلامت کند. قابل ذکر است که کشور ایران داری پیشینه پزشکی هزاران ساله است و همواره پزشکان ایرانی در سطح بین المللی شناخته شده بودند.

۵) گردشگری طبیعت درمانی

سفر درمانی به مراکز دارای منابع و خدمات درمانی طبیعت تحت نظارت پزشک را گردشگری طبیعت‌درمانی می‌گویند. این منابع و خدمات طبیعی شامل چشمه‌های آب گرم، دریاچه نمک، آفتاب درخشان، لجن درمانی، ماساژ طبی، حمام گیاهی و همچنین محیط زیست زیبا، پاک و آرام به‌منظور بهبود بیماران پوستی، تنفسی، روماتولوژی، عضلانی و یا گذراندن دوران نقاهت بعد از درمان و عمل‌های جراحی صورت می‌گیرد. به لطف قرارگیری کشور ایران در این مکان جغرافیایی و چهارفصل بودن کشور و طبیعت متنوع ایران و اقسام منابع طبیعت‌درمانی در اختیار انسان‌ها گذاشته‌شده است. در حال حاضر در ایران مراکز بسیاری وجود دارد که در کنار منابع طبیعت‌درمانی امکانات اقامتی، رفاهی و نظارت پزشکی برای گردشگران آماده کرده‌اند تا گردشگران بتوانند در آرامش و رفاه کامل از این امکانات استفاده کنند.

۶) گردشگری تندرستی

در این نوع سفر، گردشگر به دنبال آرامش است و برای رهایی از تنش‌های زندگی روزمره و تجدیدقوا، بدون مداخله و نظارت پزشکی راهی سفر می‌شود. معمولاً این گردشگران بیماری جسمی مشخصی ندارند و بیشتر در پی بهره‌مندی از طبیعت شفابخش و دوری از شلوغی‌ها، تنش‌ها و آلودگی‌های زندگی شهری هستند. یکی از جاذبه‌ها و امتیازات ایران برای گردشگری درمانی در حوزه‌ی آب‌درمانی است. به‌طوری‌که در حال حاضر بیش از ۱۰۰۰ چشمه آب‌معدنی شناسایی‌شده است. از جاذبه‌های گردشگری تندرستی در ایران، وجود هتل‌ها و ویلای جنگلی، کویری، ساحلی و یا کوهستانی است؛ که افراد می‌توانند در آرامش کامل در محیط‌زیست پاک، تندرستی خود را بهبود بخشند.

۷) گردشگری سلامت در ایران

گردشگری سلامت در ایران باوجود کیفیت بالای خدمات درمانی و قیمت ارزان، به علت نبود تبلیغات مناسب، هنوز جایگاه مناسب خود را در بازارهای جهانی نیافته است و گام‌های نخست را طی می‌کند. درحالی‌که دولت بر اساس برنامه‌ریزی خود باید تا پایان برنامه چهارم توسعه، ۳۰ درصد از نیازهای درمانی کشور را از طریق صدور کالا، خدمات پزشکی و توریسم درمانی فراهم کند. در این میان ایران رقبای مهمی نیز در منطقه دارد ازجمله رقبای مهم ایران در حوزه خلیج‌فارس می‌توان اردن و دبی را نام برد.

۸) اهداف گردشگری سلامت در ایران

- ناباروری
- دندانپزشکی
- جراحی قلب، چشم و جراحی پلاستیک
- بیماری‌های مرتبط با سرطان
- گردشگری در آب‌های شفابخش

۹) ظرفیت‌های گردشگری ایران

ایران ازنظر جاذبه‌های گردشگری طبیعی جزء 10 کشور برتر جهان است اما از بازار توریسم و گردشگری سهم شاخصی نبرده است. این درحالی‌که است که از گردشگری سلامت به‌عنوان صنعت سوم جهان نام‌برده می‌شود و برخی از کشورها مانند اسپانیا از طریق این صنعت امرارمعاش می‌کنند. امروزه با صنعتی شدن کشورها و بروز دردهای مختلف در روح و بدن انسان‌ها و روی آوردن به عمل‌های جراحی زیبایی به‌صورت چشم‌گیر، بحث توریسم درمانی یکی از بخش‌های رو به رشد صنعت گردشگری در جهان و ایران به شمار می‌آید.

جداول مربوط به میزان ورود گردشگرانی که به کشور سفرکرده‌اند، همچنین مراکز گردشگری سلامت و بیمارستان‌های میزبان گردشگران سلامت را مشاهده می‌کنید.

وضعیت گردشگران ورودی به کشور به تفکیک ماه در سال‌های ۹۵ و ۱۳۹۶ (جدول ۱)

سال ۱۳۹۵		سال ۱۳۹۶				
ماه	تعداد	درصد	تعداد	درصد	تراز	درصد تغییرات
فروردین	۳۷۰۰۵۶	۷.۵۵	۳۷۵۴۳۸	۷.۲۴	۵۳۸۲	۱.۴۵
اردیبهشت	۴۳۷۴۶۴	۷.۹۳	۴۳۳۷۵۱	۸.۴۸	-۳۷۱۳	-۰.۸۵
خرداد	۳۷۹۷۱۳	۷.۷۵	۲۹۵۶۶۱	۵.۷۸	-۸۴۰۵۲	-۲۲۱۴
تیر	۴۲۸۲۸۹	۸.۷۴	۴۹۰۶۳۱	۹.۵۹	۶۲۳۴۳	۱۴.۵۶
مرداد	۵۶۷۹۰۰	۱۱.۵۹	۵۶۴۴۰۱	۱۱.۰۴	-۳۴۹۹	-۰.۶۲
شهریور	۵۱۸۴۳۷	۱۰.۵۸	۴۵۹۴۲۸	۸.۹۸	-۵۹۰۰۹	-۱۱.۳۸
مهر	۴۱۸۵۱۹	۸.۵۴	۳۸۷۴۲۸	۷.۵۷	-۳۱۲۷۱	-۷.۴۷
آبان	۳۹۰۸۷۹	۷.۹۸	۵۰۰۸۶۵	۹.۷۹	۱۰۹۹۸۶	۲۸.۱۴
آذر	۴۱۶۱۶۵	۸.۴۹	۳۸۵۸۵۴	۷.۵۵	-۳۰۳۱۱	-۷.۲۸
دی	۲۹۸۰۳۶	۶.۰۸	۳۷۷۰۷۱	۷.۳۷	۷۹۰۳۵	۲۶.۵۲
بهمن	۳۴۶۳۹۴	۷.۰۷	۴۱۶۹۲۴	۸.۱۵	۷۰۵۳۰	۲۰.۳۶
اسفند	۳۲۹۲۳۲	۶.۷۲	۴۲۶۲۵۲	۸.۳۴	۹۷۰۲۰	۲۹.۴۷
مجموع	۴۹۰۱۰۸۴	۱۰۰.۰۰	۵۱۱۳۵۲۴	۱۰۰.۰۰	۲۱۲۴۴۰	۴.۳۳

وضعیت گردشگران ورودی به تفکیک فصل در سال‌های ۱۳۹۵ و ۱۳۹۶ (جدول ۲)

سال ۱۳۹۵		سال ۱۳۹۶				
ماه	تعداد	درصد	تعداد	درصد	تراز	درصد تغییرات
بهار	۱۱۸۷۲۳۳	۲۴.۲۲	۱۱۰۴۸۵۰	۲۱.۶۱	-۸۲۳۸۳	-۶.۹۴
تابستان	۱۵۱۴۶۲۶	۳۰.۹۰	۱۵۱۴۴۶۰	۲۹.۶۲	-۱۶۶۰	-۰.۰۱
پاییز	۱۲۲۵۵۶۳	۲۵.۰۱	۱۲۷۳۹۶۷	۲۴.۹۱	۴۸۴۰۴	۳.۹۵
زمستان	۹۷۳۶۶۲	۱۹.۸۷	۱۲۲۰۲۴۷	۲۳.۸۶	۲۴۶۵۸۵	۲۵.۳۳
مجموع	۴۹۰۰۱۰۸۴	۱۰۰.۰۰	۵۱۱۳۵۲۴	۱۰۰.۰۰	۲۱۲۴۴۰	۴.۳۳

یازده کشور اول مبدأ گردشگر ورودی در سال‌های ۱۳۹۵ و ۱۳۹۶ (جدول ۳)

		سال ۱۳۹۵		سال ۱۳۹۶			
رتبه	ملیت	تعداد	درصد	تعداد	درصد	تراز	درصد تغییرات
۱	عراق	۱۳۹۸۲۰۱	۳۲.۸۰	۱۳۵۶۱۸۵	۲۹.۸۵	-۴۲۰۱۶	-۳.۰۱
۲	آذربایجان	۹۸۱۲۰۰	۲۳.۰۲	۹۵۹۶۵۳	۲۱.۱۲	-۲۱۵۴۷	-۲.۲۰
۳	افغانستان	۶۸۳۲۴۱	۱۶.۰۳	۸۶۰۰۷۹	۱۸.۹۳	۱۷۶۸۳۸	۲۵.۸۸
۴	ترکیه	۴۴۴۲۰۰	۱۰.۴۲	۵۷۳۸۴۹	۱۲.۶۳	۱۲۹۶۴۹	۲۹.۱۹
۵	پاکستان	۲۴۲۰۸۷	۵.۶۸	۲۷۳۲۷۳	۶.۰۱	۳۱۱۸۶	۱۲.۸۸
۶	ترکمنستان	۱۰۴۸۵۹	۲.۴۶	۱۴۹۹۴۹	۳.۳۰	۴۵۰۹۰	۴۳.۰۰

		سال ۱۳۹۵		سال ۱۳۹۶			
۷	بحرین	۹۳۸۸۵	۲.۲۰	۸۸۰۱۲	۱.۹٤	-۵۸۷۳	-٦.۲٦
۸	کویت	۷۸۹۰۹	۲.۰٦	۷۸۱٤۰	۱.۷۲	-۹۷٦۹	-۱۱.۱۱
۹	هندوستان	۷۷۵۸۵	۱.۸۲	۷۲۸۹	۱.٦۰	-٤۷۷٦	-٦.۱٦
۱۰	چین	۷٤۰۳٤	۱.۷٤	٦۸۹۵۹	۱.۵۲	-۵۰٦۵	-٦.۸٤
۱۱	آلمان	۷۵۷۳۷	۱.۷۸	٦۲۹۵۸	۱.۳۹	-۱۲۷۷۹	-۱٦.۸۷
مجموع		٤۲٦۲۹۳۸	۱۰۰.۰۰	٤۵٤۳۸۷٦	۱۰۰.۰۰	۲۸۰۹۳۸	٦.۵۹

۱۰) سیستم توصیه‌گر

سیستم‌های توصیه‌گر[۲] یا پیشنهاددهنده زیرمجموعه ایی از سامانه پالایش اطلاعات که به دنبال پیش‌بینی امتیاز یا اولویتی است که کاربر به یک آیتم (داده، اطلاعات، کالا و...) خواهد داد (سون و ژانگ ۲۰۱۸). در سال‌های اخیر سیستم‌های توصیه‌گر بسیار متداول شده و در حوزه‌های مختلفی مورداستفاده قرارگرفته‌اند. برخی از کاربردهای معروف این سیستم‌ها می‌توان به موارد زیر اشاره کرد: موسیقی، صفحات وب، اخبار، کتاب‌ها و مقالات، صنعت گردشگری، صنعت هتل داری، صنعت پزشکی، جستجو و شبکه‌های اجتماعی (عربی و همکاران ۲۰۲۰). درواقع سامانه‌های توصیه‌گر مانند یک فیلتر عمل می‌کنند، فیلتری که فقط آنچه مطلوب کاربر است را به او نشان می‌دهد که به این عمل شخصی‌سازی کردن اطلاعات می‌گویند. به‌طور کل سامانه توصیه‌گر یک سامانه پشتیبان شخصی‌سازی است که اطلاعات را به سه عامل تعیین‌کننده، سفارشی‌سازی، علاقه‌مندی و سودمندی، برای کاربران ویژه به‌وسیله تجزیه‌وتحلیل سلایق آن‌ها و محتوای آیتم‌ها می‌سنجد. سیستم توصیه‌گر ازجمله ابزارهایی است که می‌تواند کاربران را در محیط‌های الکترونیکی به سمت یافتن اطلاعات، خدمات و آیتم‌های موردنظرشان هدایت کند (نکیتیم،۲۰۰۹). سیستم‌های توصیه‌گر باقابلیت کشف علایق کاربران و پیش‌بینی اولویت آن‌ها، آیتم‌هایی که احتمال می‌رود موردتوجه کاربر باشد را از بین حجم بالای داده‌ها پالایش کرده و یا آن‌ها را پیشنهاد آن‌ها، در زمان او صرفه‌جویی می‌کند. از طرف دیگر این سیستم‌ها توانایی ذخیره و تحلیل رفتارهای گذشته کاربر، خدمات و اطلاعاتی را که موردتوجه کاربران نبوده و احتمالاً به آن‌ها علاقه‌مند هستند را نیز استنتاج کرده و نتایج جالب‌توجهی به کاربران ارائه می‌کند. درواقع سیستم‌های توصیه‌گر یکی از ابزارهای اصلی غلبه بر مشکل افزونگی اطلاعات بوده و با داشتن قدرت تحلیل رفتارهای کاربر، مکمل هوشمندی برای مفاهیم بازیابی و پالایش اطلاعات است. امروزه سیستم‌های توصیه‌گر در زمینه‌های متنوعی از پالایش اطلاعات موجود در وب متناسب با خواسته‌های کاربر تا تجارت الکترونیکی، پیشنهاد فیلم، موزیک، کتاب، مقاله و... کاربرد دارند (بوبادیا و همکاران،2013؛ وانگ و همکاران،2015).

۱۱) داده‌کاوی

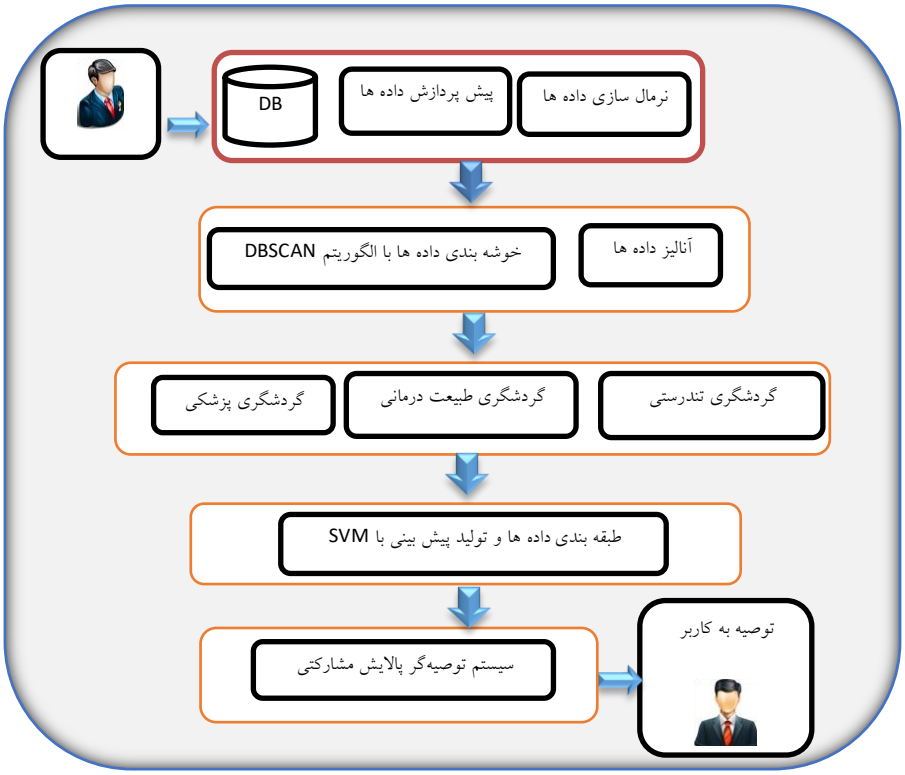
مرکز تحقیقات آمریکا و اداره پاسخگویی سازمان‌ها داده‌کاوی را مستلزم استفاده از ابزارهای پیشرفته برای تحلیل و کشف روابط و استخراج الگوهای ارزشمند داده‌ها به‌منظور دسترسی به قوانین جدید معنادار می‌داند. کشف دانش درون داده‌ها در عصر اطلاعات یکی از هیجان‌انگیزترین و کلیدی‌ترین مفاهیمی است که روزبه‌روز به اهمیت آن افزوده می‌شود (مرادی

[2] Recommender system

و همکاران،2013). با توجه به تعاریف و تفاسیر مطرح‌شده از دیدگاه‌های مختلف، می‌توان دو جزء اساسی را در داده‌کاوی مشخص نمود، اولی کشف الگوهای پنهان در داده‌ها است و دیگری استفاده از این الگوها برای پیش‌بینی نتایج در آینده است. در سال‌های اخیر با پیشرفت سریع و گسترده شبکه‌ها و فناوری پایگاه‌های اطلاعاتی و داده‌ها هم با حجم بسیار زیادی از اطلاعات مواجه بوده‌اند. (پناهی،2018)، (وو و همکاران،2013؛ رویگر،2017).

۱۲) روش پیشنهادی

در روش پیشنهادی به ارائه روشی جهت بهبود سیستم‌های توصیه‌گر در زمینه گردشگری پرداخته می‌شود، بدین‌صورت که ابتدا باید اطلاعات مربوط به تمام مناطق گردشگری سلامت و تمامی پزشکان و متخصصان این زمینه جمع‌آوری شود. سپس باید عملیات پیش‌پردازش داده‌ها را بر روی پایگاه داده موجود انجام دهیم زیرا نمی‌توان داده‌ها را به‌صورت خام به الگوریتم‌های داده‌کاوی تزریق کرد. در مرحله بعد باید داده‌ها را به‌منظور پاک کردن داده‌های اضافی و بدون استفاده پاک‌سازی کنیم. پس از عملیات پاک‌سازی داده‌ها باید آن‌ها را نرمال‌سازی کنیم، هدف نرمال‌سازی حذف افزونگی و باقی نگه‌داشتن وابستگی بین داده‌هاست. در مرحله بعد باید داده ها را خوشه بندی کنیم. برای خوشه بندی داده ها از روش خوشه‌بندی DBSCAN استفاده خواهیم کرد، سپس شباهت داده‌ها را با استفاده از الگوریتم اقلیدسی محاسبه می‌کنیم. سپس با استفاده از الگوریتم SVM داده‌ها را دسته‌بندی می‌کنیم سپس در مرحله آخر با استفاده از سیستم توصیه‌گر مبتنی برپالایش مشارکتی داده‌هایی به کاربر پیشنهاد داده می‌شود که می‌تواند موردعلاقه گردشگران باشند. در شکل زیر نمایی از روش پیشنهادی موردنظر را مشاهده می‌کنید (شکل ۱).



نمایی از روش پیشنهادی موردنظر (شکل ۱)

۱۳) پایگاه داده

در ابتدا نیاز داریم که لیستی از اماکن گردشگری سلامت و متخصصان مربوطه در این زمینه را جمع‌آوری کنیم. لیست مربوطه حاوی نقاط مربوط به اماکن گردشگری و متخصصان مربوطه در حوزه سلامت است که باید شامل آدرس دقیق مکان‌های موردنظر باشد. سپس در مرحله بعد باید این لیست را مطابق روش پیشنهادی مورد پیش‌پردازش و آنالیز قرار دهیم.

۱۴) پیش‌پردازش داده‌ها

در مرحله اول روش پیشنهادی ابتدا باید عمل پیش‌پردازش داده‌ها را انجام دهیم در فرآیند داده‌کاوی مانند طبقه‌بندی و خوشه‌بندی نیاز داریم تا داده‌ها برای الگوریتم آماده شوند، زیرا معمولاً نمی‌توان داده‌ها را به‌صورت خام به الگوریتم‌های داده‌کاوی و یادگیری ماشین تزریق کرد. برای آماده‌سازی داده‌ها، نیاز است تا آن‌ها را از شکل و حالت اولیه، خارج کرده و به شکلی که برای الگوریتم مناسب باشد تبدیل کرد. همچنین داده‌های موجود معمولاً دارای زوائد مختلفی هستند که ممکن است الگوریتم را دچار خطا کنند. در داده‌کاوی هم نیاز داریم تا داده‌های اضافی که به مسئله و الگوریتم کمکی نمی‌کنند را حذف کنیم. عملیات پیش‌پردازش داده‌ها معمولاً قبل از عملیات اصلی الگوریتم‌های داده‌کاوی انجام می‌گیرند و باعث تسهیل و کمک به الگوریتم‌ها می‌شوند.

پس از پیش‌پردازش داده‌ها باید داده‌ها را نرمال کنیم، نرمال‌سازی داده‌ها تغییر داده‌ها به گونه ایی است که آن‌ها را به یک دامنه کوچک و معین مانند فاصله بین ۱- و ۱ نگاشت کنند. هدف نرمال‌سازی حذف افزونگی داده‌ها و باقی نگه‌داشتن وابستگی بین داده‌های مرتبط است. این فرآیند اغلب باعث ایجاد جداول بیشتر می‌شود ولی اندازه‌گیری پایگاه داده را کاهش داده و بهبود کارایی را تضمین می‌کند. روش‌های مختلفی جهت نرمال‌سازی داده‌ها وجود دارد که معروف‌ترین آن‌ها می‌توان به روش Min − Max Normalization اشاره کرد. در این روش هرکدام از داده‌ها را می‌توان به یک بازه دلخواه تبدیل کرد. فرمول کلی این روش برای تبدیل داده‌ها به بازه بین ۰ تا ۱ به‌صورت زیر هست.

رابطه ۱)

$$Z = \frac{X - MIN(X)}{MAX(X) - MIN}$$

۱۵) خوشه‌بندی داده‌ها

در این مرحله باید داده‌ها را خوشه‌بندی کنیم. روش خوشه‌بندی موردنظر الگوریتم خوشه‌بندی DBSCAN است. این الگوریتم یکی از رایج‌ترین الگوریتم‌های استفاده‌شده در سیستم‌های توصیه‌گر است. به‌عنوان معیار شباهت نیز از روش اقلیدسی استفاده می‌شود. درواقع در این بخش تعداد داده‌ها خوشه‌بندی می‌شود و برای هر خوشه یک نماینده و یا یک بردار میانگین در نظر گرفته می‌شود. نحوه کار این الگوریتم بدین‌صورت است که DBSCAN با یک نقطه اولیه دلخواه شروع می‌شود که بازدید نشده است. محدوده این نقطه با استفاده از فاصله epsilon استخراج می‌شود (تمام نقاطی که در فاصله ε هستند نقاط هم‌گروه یا همسایه هستند). باید به یادداشت که الگوریتم برای پیدا کردن همسایگی در یک فضای دوبعدی و سه‌بعدی از فاصله اقلیدسی استفاده می‌کند به‌این‌ترتیب همسایگی توسط کمترین مقدار فاصله از نقطه اصلی

تعریف می‌شود. اگر تعداد کافی نقاط [3] در این محدوده وجود داشته باشد، فرآیند خوشه سازی شروع می‌شود و نقطه داده فعلی به اولین نقطه خوشه در خوشه جدید تبدیل می‌شود و در غیر این صورت، نقطه به‌عنوان نویز تلقی می‌شود (بعدها این نقطه نویز ممکن است بخشی از خوشه شود). در هر دو مورد این نقطه به‌عنوان بازدید شده مشخص می‌شود. برای این اولین نقطه در خوشه جدید، نقاط در محدوده ε فاصله آن نیز بخشی از یک خوشه است. این روش برای ساخت همه نقاط در گروه ε متعلق به یک خوشه مشابه است و سپس برای همه نقاط جدید که فقط به گروه خوشه اضافه‌شده‌اند تکرار می‌شود. این فرآیند تکرار می‌شود تا تمام نقاط در خوشه‌ها وارد شوند یعنی همه نقاط در محدوده ε خوشه‌ای بازدید شده و برچسب‌گذاری شده‌اند (تنی 2018)(سیدرات، 2019)

۱۶) محاسبه شباهت خوشه ها

خوشه‌بندی فرآیند خودکارسازی است که در طی آن، اشیاء به دسته‌هایی که اعضای آن‌ها ازنظر شاخص‌های موردنظر مشابه یکدیگر باشند تقسیم می‌شوند؛ بنابراین برای سنجش شباهت بین اشیای داده از اندازه‌گیری فاصله استفاده می‌شود. روش‌های مختلفی برای اندازه‌گیری فاصله بین دو شی وجود دارد که فاصله اقلیدسی معروف‌ترین و پرکاربردترین گونه فاصله است که به‌صورت رابطه زیر محاسبه می‌شود.

رابطه ۲)

$$d(x,y) = \sqrt{\sum_{i=1}^{m} (xi - yi)^2}$$

۱۷) آنالیز داده‌ها

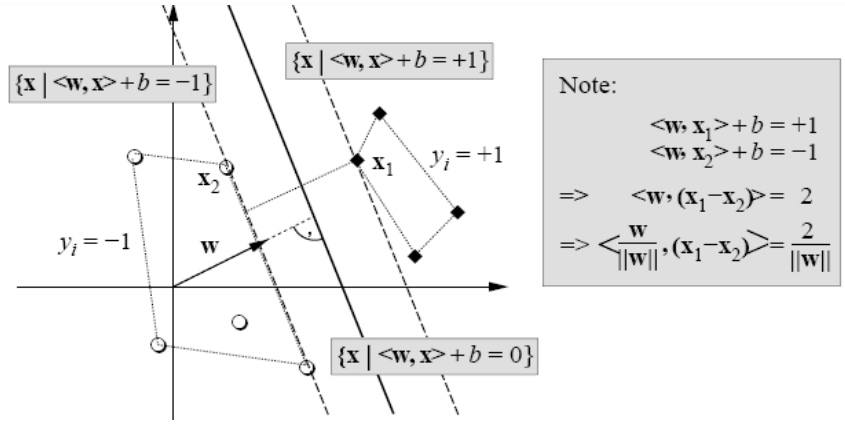
در این بخش داده‌های خوشه‌بندی‌شده از مرحله قبل را بر اساس هدف گردشگر جهت ورود به کشور مورد آنالیز قرار می‌دهیم و به سه بخش اصلی گردشگری پزشکی، گردشگری طبیعت‌درمانی و گردشگری تندرستی تقسیم‌بندی می‌کنیم. سپس کاربر با انتخاب هرکدام از این سه آیتم هدف سفر خود را مشخص خواهد کرد. در صورت انتخاب گردشگری پزشکی، لیستی از متخصصین کشور بر اساس نوع تخصص به وی معرفی خواهد شد و در صورت هر یک از دو گزینه گردشگری طبیعت‌درمانی یا گردشگری تندرستی، تعدادی لوکیشن از مناطق مختلف کشور که دارای اماکن گردشگری سلامت هستند از طریق سیستم توصیه‌گر مبتنی بر فیلترینگ مشارکتی به کاربر پیشنهاد خواهد شد که می‌تواند موردعلاقه وی باشند.

۱۸) روش طبقه‌بندی ماشین بردار پشتیبان

در مرحله بعد جهت دسته‌بندی و پیش‌بینی دقیق‌تر داده‌ها از SVM استفاده می‌کنیم. فرض کنیم مجموعه نقاط داده $\{(x_1, c_1), (x_2, c_2), \dots, (x_n . c_n)\}$را در اختیارداریم و می‌خواهیم آن‌ها را به دوطبقه $c_i = \{-1,1\}$ تفکیک کنیم. هر x_i یک بردار p بعدی از اعداد حقیقی است که درواقع همان متغیرهای بیانگر رفتار سیستم هستند.

[3] Min points

روش‌های طبقه‌بندی خطی، سعی دارند که با ساختن یک ابر سطح (که عبارت است از یک معادله خطی)، داده‌ها را از هم تفکیک کنند. روش طبقه‌بندی ماشین بردار پشتیبان که یکی از روش‌های طبقه‌بندی خطی است، بهترین ابر سطحی را پیدا می‌کند که با **حداکثر فاصله**، داده‌های مربوط به دوطبقه را از هم تفکیک کند. به‌منظور درک بهتر مطلب، در شکل ۲ تصویری از یک مجموعه داده متعلق به دو کلاس نشان داده‌شده که روش ماشین بردار پشتیبان بهترین ابر سطح را برای جداسازی آن‌ها انتخاب می‌کند.



نحوه ساخت ابر سطح جداکننده بین دوطبقه داده در فضای دوبعدی شکل (۲)

۱۸-۱) نحوه حل مسئله در حالت کلی

مسئله، به حداقل رساندن تابع $f(x)$ مشروط به محدودیتِ $g(x) = 0$ است. شرط لازم برای اینکه x_0 یک جواب باشد به شرح زیر است:

رابطه ۳)
$$\begin{cases} \frac{\partial}{\partial x}\left(f(x) + \alpha g(x)\right)|_{x=x_0} = 0 \\ g(x) = 0 \end{cases}$$

α ضریب لاگرانژ[4] است. برای چندین محدودیت $g_i(x) = 0$ ، i=1,…,m ، نیاز به ضریب لاگرانژ α_i برای هر یک از محدودیت‌ها است. آنگاه رابطه (۶) به شکل زیر نمایش داده می‌شود.

رابطه ٤)
$$\begin{cases} \frac{\partial}{\partial x}\left(f(x) + \sum_{i=1}^{n} \alpha_i g_i(x)\right)|_{x=x_0} = 0 \\ g_i(x) = 0 \qquad \text{for } i = 1, \dots, m \end{cases}$$

در حالتی که محدودیت به‌صورت یک نامساوی باشد، $g_i(x) \le 0$ بازهم مسئله به همان صورت است تنها با این تفاوت که α_i باید مثبت باشد. در این حالت اگر x_0 جوابی برای مسئله بهینه‌سازی باشد، آنگاه باید برای هر $i = 1, \dots, m$ وجود داشته باشد $\alpha_i \ge 0$ به‌گونه‌ای که x_0 در شرایط (۸) صدق کند.

رابطه ۵)
$$\begin{cases} \frac{\partial}{\partial x}\left(f(x) + \sum_{i=1}^{n} \alpha_i g_i(x)\right)|_{x=jx_0} = 0 \\ g_i(x) \le 0 \qquad \text{for } i = 1, \dots, m \end{cases}$$

تابع $f(x) + \sum_{i=1}^{n} \alpha_i g_i(x)$ تابع لاگرانژی نام دارد. شیب این تابع باید مساوی صفر قرار داده شود.

[4] Lagrange multiplier

۱۸-۲) فیلترینگ مشارکتی

پالایش مشارکتی فرآیند پالایش یا ارزیابی آیتم‌ها با استفاده از نظرات کاربران است. باوجوداینکه از شکل‌گیری اصطلاح پالایش مشارکتی تنها کمی بیش از یک دهه می‌گذرد، اما فلسفه این روش این است که استفاده از نظرات دیگران در تصمیم‌گیری است، برای قرن‌ها مورداستفاده‌ی انسان‌ها بوده است. برای مثال اگر دوستان شما از مکانی تعریف کنند راغب به مراجعه به آن مکان خواهید شد. یا برعکس اگر مکانی را بد توصیف کنید بعید است برای رفتن به آن اقدام کنید. بعلاوه بعد از مدتی شما خواهید فهمید که نظرات کدام دوستانتان به نظرات شما نزدیک‌تر است و به‌تدریج فقط به نظرات آن دسته از دوستانتان که به شما شباهت دارند توجه خواهید کرد. کامپیوتر و بستر وب این امکان را فراهم کردند که قدم از ارتباطات شخصی فراتر گذاریم و به‌جای تصمیم‌گیری بر مبنای ده یا صدنفر از دوستان و نزدیکان، از نظرات بیش از میلیون‌ها کاربر استفاده کنیم. سرعت کامپیوتر این اجازه را می‌دهد که این نظرات را به‌صورت بلادرنگ پردازش کرده و بدانیم که افرادی که به ما شبیه هستند در مورد یک آیتم خاص که از آن بی‌اطلاعیم چه نظری دارند.

۱۸-۳) نحوه امتیازدهی

در این قسمت به مفهوم امتیاز و ماتریس امتیاز پرداخته می‌شود. ازآنجایی‌که در تکنیک پالایش مشارکتی نظرات کاربران نقش کلیدی دارد لازم است روش‌ها و قالب‌هایی برای جمع‌آوری آن طراحی شود. در ادبیات سیستم‌های توصیه‌گر روش‌های مختلفی برای جمع‌آوری نظرات کاربران معرفی شدند، اما روش معمولی که اکثر سیستم‌های توصیه‌گر مبتنی بر پالایش به کاربر گرفته می‌شود در نظر گرفتن یک بازه عددی (مثلاً ۱ تا ۵) برای هر آیتم، تعریف مفهوم هر یک از این اعداد (مثلاً ۱: بسیار بد، ۲: بد، ۳: متوسط، ۴: خوب، ۵: بسیار خوب) و درخواست از کاربر برای نگاشت یکی از این اعداد به هر یک از آیتم‌هایی که مشاهده می‌کند. این اعداد در ادبیات سیستم‌های توصیه‌گر امتیاز خوانده می‌شود و این روش امتیازدهی نام دارد.

در جدول زیر ماتریس امتیاز یک سیستم توصیه‌گر مبتنی بر فیلترینگ مشارکتی به چند پزشک را مشاهده می‌کنید.

امتیازدهی بدین‌صورت هست که به‌عنوان‌مثال کاربر ۳ دکتر فلاح‌پور ۳ و کاربر ۱ به دکتر علیپور امتیاز نداده است (جدول ٤)

شناسه کاربر	دکتر علیپور	دکتر شجاعی	دکتر علی عبدی	دکتر فلاح‌پور
کاربر ۱		۵	۲	۳
کاربر ۲	٤		۳	
کاربر ۳			۵	۳
کاربر ٤	۵	۵		٤

با استفاده از سیستم توصیه‌گر مبتنی بر پالایش مشارکتی مکان‌هایی را به کاربر توصیه می‌کند که می‌تواند موردعلاقه کاربر باشد. برای پیش‌بینی اماکن، از تکنیک‌های CF انجام می‌شود، تکنیک های CF از یک دیتابیس از ترجیحات برای آیتم‌های توسط کاربران استفاده می‌کنند که برای تخمین موضوعات اضافی یا ایجاد کاربران جدید است. در یک سناریوی معمولی CF، لیستی از m کاربر وجود دارد $\{u1, u2, u3..um\}$و همچنین یک لیست از n آیتم $\{i1, i2, i3...in\}$و هر

کاربر Ui، یک لیست از آیتم‌ها دارد lui که کاربر آن‌ها را رتبه‌بندی کرده است و یا آن‌هایی که ترجیحات آن‌ها از طریق رفتارشان استنباط کرده است. رتبه می‌تواند اشاره‌های صریح باشد و... که روی مقیاس ۱ تا ۵ است و یا همچنین می‌تواند اشاره ضمنی باشد.

۱۹) ارزیابی سیستم پیشنهادی

اغلب برای اعتبارسنجی سیستم‌های توصیه‌گر به مواردی همچون کارایی روش خوشه‌بندی و دقت سیستم و فراخوانی استفاده می‌شود. در این تحقیق نیز از این معیارها برای ارزیابی سیستم استفاده گردیده است. دقت و فراخوانی در سیستم‌های توصیه‌گر با استفاده از دو رابطه زیر محاسبه می‌شود.

دقت با استفاده از رابطه زیر محاسبه می‌گردد:

رابطه ۶)
$$\boldsymbol{Precision} = \frac{|\{\boldsymbol{relevant\ pages}\} \cap \{\boldsymbol{retrved\ pages}\}|}{|\{\boldsymbol{retrived\ pages}\}|}$$

فراخوانی با استفاده از رابطه زیر محاسبه می‌گردد:

رابطه ۷)
$$\boldsymbol{Recall} = \frac{|\{\boldsymbol{relevant\ peges}\} \cap \{\boldsymbol{retrved\ pages}\}|}{|\{\boldsymbol{relevant\ pages}\}|}$$

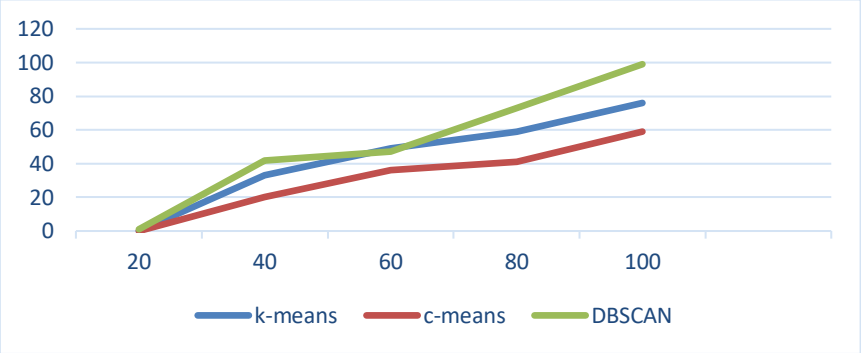
۲۰) ارزیابی کارایی الگوریتم DBSCAN

برای مقایسه کارایی خوشه‌بندی روش پیشنهادی از شبه کد پیاده‌سازی شده زیر با استفاده از زبان برنامه‌نویسی Peyton استفاده‌شده است و مقایسه ایی بین الگوریتم پیشنهادی موردنظر و الگوریتم خوشه‌بندی K-Means صورت گرفت که به‌صورت زیر است.

```python
from sklearn.cluster import DBSCAN

from sklearn.preprocessing import StandardScaler

scaler = StandardScaler()

X_scaled = scaler.fit_transform(X)
# cluster the data into five clusters

dbscan = DBSCAN(eps=0.123, min_samples = 2)

clusters = dbscan.fit_predict(X_scaled)
# plot the cluster assignments

plt.scatter(X[:, 0], X[:, 1], c=clusters, cmap="plasma")

plt.xlabel("Feature 0")

plt.ylabel("Feature 1")
```

پیاده‌سازی الگوریتم خوشه‌بندی DBSCAN در پایتون (شکل ٤)

نتایج حاصل از شبه کد بالا حاکی از آن است که کارایی الگوریتم خوشه‌بندی پیشنهادی حدود ۰/۹۶ است این در حالی است که الگوریتم خوشه‌بندی K-MEANS امتیاز کارایی ۰/۷۶ را به دست آورد که نشان از کارایی الگوریتم DBSCAN نسبت به سایر الگوریتم‌های خوشه‌بندی است.

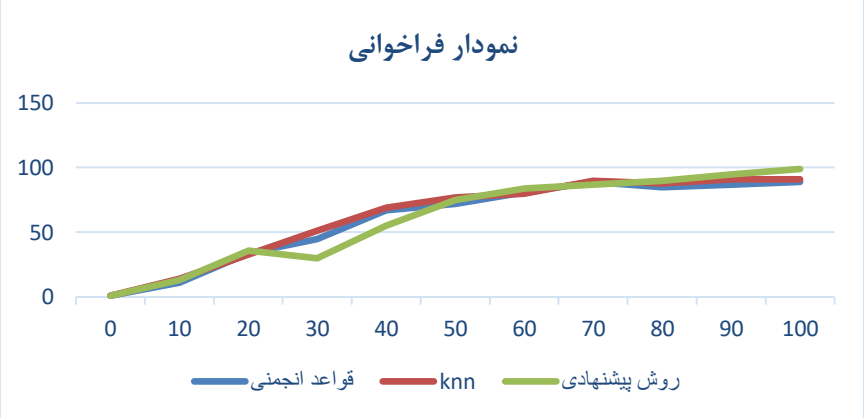


نمودار برسی عملکرد میزان کارایی الگوریتم DBSCAN و سایر الگوریتم‌های خوشه‌بندی (شکل ٦)

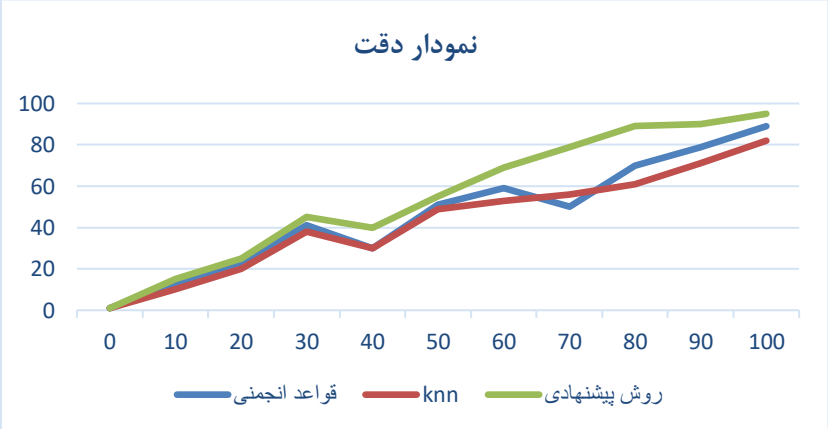
۲۱) ارزیابی میزان دقت و فراخوانی سیستم پیشنهادی

جدول ارزیابی میزان دقت و فراخوانی روش پیشنهادی (جدول ٥)

نام الگوریتم	KNN	قواعد انجمنی	روش پیشنهادی
دقت	۸۲٪	۸۹٪	۹۵٪
فراخوانی	۸۹٪	۹۱٪	۹۹٪



مقایسه فراخوانی مربوط به روش پیشنهادی و سایر الگوریتم‌ها (شکل ۷)



مقایسه دقت مربوط به روش پیشنهادی و سایر الگوریتم‌ها (شکل ۸)

به‌عنوان‌مثال برای خروجی سیستم فرض می‌کنیم در جدول زیر تعداد ۱۰ گردشگر داریم که می‌خواهند برای مناطق گردشگری سلامت کشور و ورود به کشور اقدام نمایند.

کشور	تعداد
ترکیه	۳
سنگاپور	۱
تایلند	۱
عراق	۵

مثلاً ۲ کاربر از کشور ترکیه قصد سفر به چشمه‌های آب گرم را دارند و ۲ نفر از گردشگران عراقی قصد درمان خود را در یکی از مراکز سلامت کشور دارند. پس از مراجعه این گردشگران به سامانه توصیه‌گر پیشنهادی از کاربر درخواست می‌شود که هدف از سفر خود به کشور را با انتخاب یکی از ۳ گزینه موجود در روش پیشنهادی شامل گردشگری تندرستی، گردشگری طبیعت‌درمانی و گردشگری پزشکی، انتخاب نماید. به‌طور مثال گردشگری که از کشور ترکیه قصد ورود به کشور را دارد و در نظر دارد که به چشمه‌های آب گرم دسترسی پیدا کند، با انتخاب قسمت گردشگر طبیعت‌درمانی، فهرستی از چشمه‌های آب گرم کشور در اختیار وی قرار می‌گیرد که این لیست بر اساس امتیازی سایر گردشگران ورودی که سابقه سفر به مناطقی که این چشمه‌ها در آنجا قرار دارند و امتیاز بیشتری به این مناطق داده‌اند مرتب خواهد شد و در اختیار گردشگر هدف قرار خواهد گرفت. یا گردشگر دیگر که از کشور عراق برای درمان خود قصد سفر به کشور را دارد با انتخاب گزینه گردشگر پزشکی، سیستم پیشنهادی با آنالیز پایگاه داده خود فهرستی از پزشکان متخصص مرتبط با بیماری گردشگر و نام مراکزی که این پزشکان در آن مرکز مشغول به طبابت هستند در اختیار گردشگر هدف قرار خواهد گرفت و این گردشگران می‌توانند با حضور در این مراکز خود را به دستان این متخصصان بسپارند و مراحل درمان خود را با خیالی آسوده سپری نمایند.

۲۲) بحث و نتیجه‌گیری

در این تحقیق به ارائه روش جدیدی به‌منظور بهبود سیستم‌های توصیه‌گر در زمینه گردشگری سلامت پرداخته شد. روش کار این سیستم بدین‌صورت است که پس از درخواست گردشگر موردنظر، با استفاده از اطلاعات و امتیازاتی که سایر گردشگرانی که قبلاً به کشور سفرکرده‌اند به اماکن و متخصصان کشور داده‌اند لیستی از این اماکن در اختیار گردشگر قرار می‌گیرد که می‌تواند موردعلاقه وی باشد، در این سیستم ابتدا اطلاعات مربوط به کاربران جمع‌آوری و مورد آنالیز قرار می‌گیرند سپس این اطلاعات را با استفاده از الگوریتم DBSCAN خوشه‌بندی می‌کنیم سپس اطلاعات خوشه‌بندی‌شده را با استفاده از SVM دسته‌بندی کردیم، سپس با استفاده از سیستم توصیه‌گر مبتنی بر فیلترینگ مشارکتی و امتیازات کاربران، اطلاعات مفید و دارای بیشترین امتیاز را در اختیار کاربر هدف قرار می‌دهیم. نحوه کار سیستم بدین‌صورت است که پس از درخواست کاربر از سیستم موردنظر از کاربر درخواست می‌شود که نوع سفر خود را با استفاده از سه گزینه گردشگری پزشکی، گردشگری تندرستی و گردشگری طبیعت‌درمانی انتخاب کند که پس از انتخاب یک مورد از گزینه‌ها فهرستی از اماکن در اختیار کاربر قرار خواهد گرفت که می‌تواند موردعلاقه وی باشد. نتایج این تحقیق حاکی از ارزیابی

و مقایسه مدل پیشنهادی با روش‌های دیگر حاکی دقت ۸۰ درصدی این سیستم توصیه‌گر هست که این ارزیابی از طریق دو رابطه دقت و فراخوانی مورد ارزیابی قرارگرفته شده است و این سیستم می‌تواند پیش‌بینی‌های دقیقی داشته باشد. ازآنجایی که صنعت گردشگری سلامت در آینده ایی نزدیک به رتبه اول درآمد اقتصادی جهان تبدیل خواهد شد و ایران با دارا بودن مکان‌های گردشگری پرشمار و همچنین با دارا بودن پزشکان متبحر در تمام زمینه‌های درمانی همچنین خاستگاه غنی و ارزشمند ازلحاظ طبیعی و توریستی، می‌تواند یکی از بهترین راه‌ها برای کسب درآمد برای کشور باشد. گردشگری پزشکی در ایران باوجود کیفیت بالای خدمات درمانی و قیمت ارزان، به علت نبود تبلیغات مناسب، هنوز جایگاه خود را آن‌طور که سزاوار است نیافته. ایران ازلحاظ جاذبه‌های گردشگری طبیعی جزء ۱۰ کشور برتر جهان است ولی متأسفانه از بازار توریسم و گردشگری، سهم قابل‌توجهی نبرده است. ایران توانایی جذب سالانه حدود ۱ میلیون گردشگر پزشکی را دارد که برابر با درآمدی معادل ۷ میلیارد دلار برای کشور خواهد بود. ازاین‌رو وجود سیستمی که بتواند مکان‌های توریستی و گردشگری سلامت را به کاربران خارجی معرفی کند لازم و ضروری به نظر می‌رسد، لذا در این تحقیق سعی شد به روشی جهت این کار ارائه شود که گردشگران را راغب به سفر به کشور زیبای ایران کند تا هم بتواند گردشگران سلامت را درمان و همچنین آن‌ها را بافرهنگ اصیل ایرانی آشنا کند.

منابع

Arabi, H., Balakrishnan, V., & Shuib, N. L. M. (2020). A Context-Aware Personalized Hybrid Book Recommender System. Journal of Web Engineering, 405-428. DOI: https://doi.org/10.13052/jwe1540-9589.19343

Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. Knowledge-based systems, 46, 109-132. DOI: https://doi.org/10.1016/j.knosys.2013.03.012

El Bahi, H., & Zatni, A. (2018). DOCUMENT TEXT DETECTION IN VIDEO FRAMES ACQUIRED BY A SMARTPHONE BASED ON LINE SEGMENT DETECTOR AND DBSCAN CLUSTERING. Journal of Engineering Science and Technology, 13(2), 540-557. DOI: https://doi.org/10.1007/978-3-030-12048-1_7

Gök, A., Waterworth, A., & Shapira, P. (2015). Use of web mining in studying innovation. Scientometrics, 102(1), 653-671 DOI: https://doi.org/10.1007/s11192-014-1434-0

Guha, S., & Mishra, N. (2016). Clustering data streams. In Data stream management (pp. 169-187). Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-540-28608-0_8

Hanefeld, J., Lunt, N., Smith, R., & Horsfall, D. (2015). Why do medical tourists travel to where they do? The role of networks in determining medical travel. Social Science & Medicine, 124, 356-363. DOI: https://doi.org/10.1016/j.socscimed.2014.05.016

Kagita, V. R., Pujari, A. K., & Padmanabhan, V. (2015). Virtual user approach for group recommender systems using precedence relations. Information Sciences, 294, 15-30. DOI: https://doi.org/10.1016/j.ins.2014.08.072

Ładyżyński, P., & Grzegorzewski, P. (2015). Vague preferences in recommender systems. Expert Systems with Applications, 42(24), 9402-9411. DOI: https://doi.org/10.1016/j.eswa.2015.08.006

Lu, J., Wu, D., Mao, M., Wang, W., & Zhang, G. (2015). Recommender system application developments: a survey. Decision Support Systems, 74, 12-32. DOI: https://doi.org/10.1016/j.dss.2015.03.008

Nguyen, T. N., & Ricci, F. (2018). A chat-based group recommender system for tourism. Information Technology & Tourism, 18(1-4), 5-28. DOI: https://doi.org/10.1007/s40558-017-0099-y

Nilashi, M., Bagherifard, K., Rahmani, M., & Rafe, V. (2017). A recommender system for tourism industry using cluster ensemble and prediction machine learning techniques. Computers & industrial engineering, 109, 357-368. DOI: https://doi.org/10.1016/j.cie.2017.05.016

Nkweteyim, D. (2009). Hyperlink Recommender Systems Design: A Research Study on Tools and Techniques. VDM Verlag DOI: https://doi.org/10.1007/978-0-387-85820-3_1

Rahimi, Haji Zeynalabedini ., Applications Of Datamining in science information and Epistemology.Nov 22:2(3):23:-32 2015 DOI: https://doi.org/10.5120/ijca2016909249

Roiger, R. J. (2017). Data mining: a tutorial-based primer. CRC press DOI: https://doi.org/10.1201/9781315382586

Siddharth Agrawal.)2019 (Machine learning-DBSCAN. Toward Data Science. Dor: 20.1001.1.27170411.1402.15.57.6.1

Sinha, A. K., Raj, N., Haque, S., Haque, A., & Singh, N. K. Web Content Mining: Tool, Technique & Concept DOI: https://doi.org/10.9790/0661-1806065760

Sun, Y., & Zhang, Y. (2018, June). Conversational recommender system. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (pp. 235-244). DOI: https://doi.org/10.1145/3209978.3210002

Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2013). Data mining with big data. IEEE transactions on knowledge and data engineering, 26(1), 97-107. DOI: https://doi.org/10.1109/TKDE.2013.109

Zhang, F., Yuan, N. J., Lian, D., Xie, X., & Ma, W. Y. (2016, August). Collaborative knowledge base embedding for recommender systems. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 353-362) DOI: https://doi.org/10.1145/2939672.2939673

Zhao, Y., & Lin, H. (2014, August). WEB data mining applications in e-commerce. In 2014 9th International Conference on Computer Science & Education (pp. 557-559). IEEE DOI: https://doi.org/10.1109/ICCSE.2014.6926523

Xu, Z., Chen, L., & Chen, G. (2015). Topic based context-aware travel recommendation method exploiting geotagged photos. Neurocomputing, 155, 99-107. DOI: https://doi.org/10.1016/j.neucom.2014.12.043