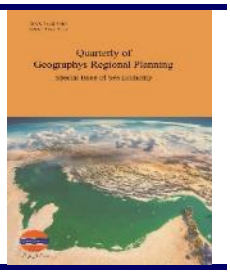
Iranian Association of Geopolitical

Geography (Regional Planning)
Autumn 2025. Vol 15. Issue 60
ISSN (Print): 2228-6462 - ISSN (Online): 2783-2112
Journal Homepage: https://www.jgeoqeshm.ir/



Research Paper

Development of an Intelligent Recommender System with a group Refinement approach based on Geographic Data Modeling

Tahereh Kheradmandan[1], Farhad Mardukhi[*2], Amir Khiekh Ahmadi[3]

1. PhD student, Computer Science Department, Sanandaj Branch, Islamic Azad University, Sanandaj, Iran
2. Department of Computer Engineering and Information Technology, Razi University of Kermanshah, Kermanshah, Iran
3. Department of Computer Engineering, Sanandaj Branch, Islamic Azad University, Sanandaj, Iran

ARTICLEINFO

PP: 464-479

Use your device to scan and read the article online

Keywords:
Development, Intelligent Recommender System, group Refinement approach, Geographic Data Modeling.

Abstract

The Main objective of this research is to design and implement an intelligent recommender system using GIS modeling data to improve the quality of recommendations in online sales systems. The selected research method is the scientific design and experimental research method, in which, in the evaluation stage of the scientific design research method, experimental design has been used. To improve the performance of the intelligent recommender system, user-defined labels and deep neural network algorithms along with GIS modeling have been used to generate recommendations. In the designed and conducted experiment, the label-based recommender system (which is designed and created to generate recommendations) is compared with the group filtering recommender system (which is a conventional and benchmark recommender system) in the evaluation criteria of precision, recall, and F1. Based on the results, the proposed recommender system based on the group filtering method and GIS data based on deep neural networks performs better than the group filtering system in all these evaluation criteria.

Citation: Kheradmandan, T., Mardoukhi, F., & Sheikh Ahmadi, A. (2025). **Development of an intelligent recommender system using a collaborative filtering approach based on geospatial data modeling**. Geography (Regional Planning Quarterly), 15(60).464-479

DOI: 10.22034/jgeoq.2025.551855.4347

[*]**Corresponding author**: Farhad Mardukhi, **Email:** mardukhi@razi.ac.ir

Copyright © 2024 The Authors. Published by Qeshm Institute. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

Extended Abstract

Introduction

In general, recommender systems directly help users find content, goods, or services (such as books, digital products, movies, music, TV shows, etc.) by collecting and analyzing other users' recommendations . These systems have become an important research field since the first paper on group refinement recommender systems was written in the early 1990s, but more research is needed in the field of recommender systems, as this field is still broader and more raw than other research areas. Many researchers have added tags to traditional recommender systems to improve their performance. These tags act as a bridge between users and products, helping to solve the cold start problem. GIS modeling helps in this process by identifying spatial and geographic patterns of user behavior and matching it with product preferences. Many researchers have added tags to traditional recommender systems to improve their performance. These tags act as a bridge between users and products, helping to solve the cold start problem. GIS modeling helps in this process by identifying spatial and geographic patterns of user behavior and matching it with product preferences.

Methodology

The research method used in this study is design-experimental. Considering the problem and purpose of this research and the need to create and implement a label-based recommender system, the stages of the design science research method process have been used as the main steps of the research process. This research method includes the steps of identifying the motivations and problem, defining the goals of the solution, designing and creating, concluding, evaluating, and communicating. In the identification, problem, and motivations stage, the research problem must be defined and the value of its solution must be acknowledged. In the identification phase, the problem and motivations must define the research problem and the value of its solution must be acknowledged. In the definition of solution goals phase, the goals of a solution must be derived from the problem definition and knowledge of what is possible and feasible. In the design and creation phases, the researcher creates the artifact.

Results and Discussion

The move from static momentary GIS to temporal GIS, which improves the ability to handle a variety of spatiotemporal dynamics, has attracted considerable attention in research. In this regard, this research has investigated the performance of a label-based recommender system based on GIS modeling for recommending movies. The main research question is whether a recommender system based on deep neural networks and GIS modeling performs better than a group filtering system? To answer this question, an experiment was designed in which a label-based recommender system was compared with a group filtering system in terms of precision, recall, and F1 evaluation criteria. The results showed that the label-based system outperformed the group filtering system in all three evaluation criteria for different lengths of the recommendation list (from 1 to 200). The reason for this superiority is the ability of the label-based system to overcome the sparseness of the data by using a sparse auto encoder to extract more abstract factors from the labels. These factors enable more accurate identification of users similar to the target user and provide more accurate recommendations.

Conclusion

This research has designed and evaluated a label-based recommender system based on deep neural networks and GIS modeling for movie recommendations, which is the first time that a system that was previously developed for music and website recommendations has been used in the field of movies. From a GIS modeling perspective, the

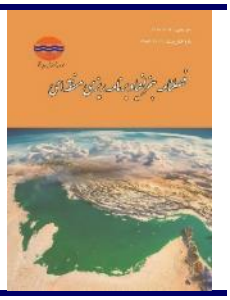
move from static momentary GIS to temporal GIS, which improves the ability to handle spatiotemporal dynamics at different scales (individual, local, mesoscale, and global) and multiple dimensions (spatial pattern, temporal trend, and statistical distribution), has played an important role in this system. The systematic combination of these dimensions and scales provides an analytical framework for identifying research gaps. The use of open data and open source toolkits such as R and Python has also facilitated stronger interdisciplinary collaboration in the development of this system. In model-driven systems, a model is built based on the data and recommendations are made based on it. These models are not updated online and are regenerated every so often, which reduces the recommendation time compared to ensemble filtering, because the number of extracted factors is less than the initial labels.

References

1. Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. Knowledge-Based Systems, 46, 109–132.
2. Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1798–1828.
3. Bengio, Y. (2009). Learning deep architectures for AI. Foundations and Trends in Machine Learning, 2(1), 1–127.
4. De Lathauwer, L., De Moor, B., & Vandewalle, J. (2000). A multilinear singular value decomposition. SIAM Journal on Matrix Analysis and Applications, 21(4), 1253–1278.
5. Firan, C. S., Nejdl, W., & Paiu, R. (2007). The benefit of using tag-based profiles. In LA-WEB 2007: Latin American Web Conference (pp. 32–41). IEEE.
6. Goodchild, M. F. (2013). Geospatial and temporal perspectives of a GIS. Annals of the American Association of Geographers, 103(5), 1072–1077.
7. Huang, B., Wu, B., & Barry, M. (2010). Geographically and temporally weighted regression for modeling spatial-temporal variation of housing prices. International Journal of Geographical Information Science, 24(3), 383–401.
8. Hotho, A., Jäschke, R., Schmitz, C., & Stumme, G. (2006). Information retrieval in folksonomies: Search and ranking. In European Semantic Web Conference (pp. 411–426). Springer.
9. Hevner, A., & Chatterjee, S. (2010). Design research in information systems: Theory and practice (Vol. 22, pp. 136–215). Springer.
10. Kim, H. N., Ji, A. T., Ha, I., & Jo, G. S. (2010). Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation. Electronic Commerce Research and Applications, 9(1), 73–83.
11. Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. Computer, 42(8), 30–37.
12. Liang, H., Xu, Y., Li, Y., Nayak, R., & Tao, X. (2010). Connecting users and items with weighted tags for personalized item recommendations. In Proceedings of the 21st ACM Conference on Hypertext and Hypermedia (pp. 51–60). ACM.
13. Liu, D. C., & Nocedal, J. (1989). On the limited memory BFGS method for large scale optimization. Mathematical Programming, 45(1–3), 503–528.
14. Motaharinejad, M., Zolfagharzadeh, M., Khadangi, E., & Sadabadi, A. (2016). Designing a model for improving banking recommender systems based on predicting customers' interests: Application of data mining techniques. Journal of Information Technology Management, 8(2), 393–314.
15. Marinho, L. B., & Schmidt-Thieme, L. (2008). Collaborative tag recommendations. In Data Analysis, Machine Learning and Applications (pp. 533–540). Springer.
16. Núñez-Valdéz, E. R., Lovelle, J. M. C., Martínez, O. S., García-Díaz, V., de Pablos, P. O., & Marín, C. E. M. (2012). Implicit feedback techniques on recommender systems applied to electronic books. Computers in Human Behavior, 28(4), 1186–1193.

17. Park, D. H., Kim, H. K., Choi, I. Y., & Kim, J. K. (2012). A literature review and classification of recommender systems research. Expert Systems with Applications, 39(11), 10059–10072.
18. Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to recommender systems handbook. In Recommender Systems Handbook (pp. 5–35). Springer.
19. Schafer, J. B., Frankowski, D., Herlocker, J., & Sen, S. (2007). Collaborative filtering recommender systems. In The Adaptive Web (pp. 291–324). Springer.
20. Su, X., & Khoshgoftaar, T. M. (2009). Survey of collaborative filtering techniques. Advances in Artificial Intelligence, 2009, Article 421425.
21. Sohrabi, B., Raeesi Vanani, I., & Zareh Mirkabad, F. (2016). Designing a recommender system for optimizing and managing bank facilities through the utilization of clustering and classification algorithms. Modern Research in Decision Making, 1(2), 53–76.
22. Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2000). Application of dimensionality reduction in recommender system: A case study (Technical Report No. TR-00-043). University of Minnesota, Department of Computer Science.
23. Ye, X., & Rey, S. J. (2013). A framework for exploratory space–time data analysis of economic data. The Annals of Regional Science, 50(1), 315–339.
24. Zhang, Z. K., Liu, C., Zhang, Y. C., & Zhou, T. (2010). Solving the cold-start problem in recommender systems with social tags. EPL (Europhysics Letters), 92(2), 28002.
25. Zuo, Y., Zeng, J., Gong, M., & Jiao, L. (2016). Tag-aware recommender systems based on deep neural networks. Neurocomputing, 204, 51–60.

فصلنامه جغرافیا (برنامه ریزی منطقه ای)
دوره ۱۵، شماره ۶۰، پاییز ۱۴۰۴
شاپا چاپی: ۶۴۶۲- ۲۲۲۸ شاپا الکترونیکی: ۲۱۱۲ -۲۷۸۳
Journal Homepage: https://www.jgeoqeshm.ir/
انجمن ژئوپلیتیک ایران



مقاله پژوهشی

توسعه سامانه توصیه‌گر هوشمند با رویکرد پالایش گروهی بر پایه مدل‌سازی داده‌های جغرافیایی

طاهره خردمندان- دانشجوی دکتری گروه کامپیوتر، واحد سنندج، دانشگاه آزاد اسلامی، سنندج، ایران
فرهاد مردوخی*- گروه مهندسی کامپیوتر و فنآوری اطلاعات، دانشگاه رازی کرمانشاه، کرمانشاه، ایران
امیر شیخ احمدی- گروه مهندسی کامپیوتر، واحد سنندج، دانشگاه آزاد اسلامی، سنندج، ایران

اطلاعات مقاله

شماره صفحات: ۴۶۴-۴۷۹

از دستگاه خود برای اسکن و خواندن مقاله به صورت آنلاین استفاده کنید

واژه‌های کلیدی:
توسعه، سامانه توصیه‌گر هوشمند، رویکرد پالایش گروهی، مدل‌سازی داده‌های جغرافیایی.

چکیده

هدف اصلی این تحقیق، طراحی و پیاده‌سازی سامانه توصیه‌گر هوشمند با استفاده از داده‌های مدل‌سازی GIS جهت بالا بردن کیفیت پیشنهادات در سیستم‌های فروش انلاین می‌باشد. روش تحقیق انتخا ب‌شده، روش تحقیق علمی طراحی و آزمایشی است که در آن، در مرحله ارزیابی روش تحقیق علمی طراحی، از طراحی آزمایش استفاده شده است. برای بهبود عملکرد سامانه توصیه‌گر هوشمند، از برچسب‌های گذاشته شده توسط کاربران و الگوریتم شبکه‌های عصبی عمیق به همراه مدل‌سازی GIS برای تولید توصیه‌ها استفاده شده است. در آزمایش طراحی و انجام شده، سامانه توصیه‌گر برچسب‌محور(که برای تولید توصیه طراحی و ایجاد شده است) با سامانه توصیه‌گر پالایش گروهی(که از سامانه‌های توصیه‌گر مرسوم و معیار است) در معیارهای ارزیابی دقت، به‌یادآوری و F1 مقایسه می‌شود. براساس نتایج، سامانه توصیه‌گر ارائه شده مبتنی بر روش پالایش گروهی و داده های GIS بر مبنای شبکه‌های عصبی عمیق در تمام این معیارهای ارزیابی از سامانه پالایش گروهی بهتر عمل می‌کند.

استناد: خردمندان، طاهره؛ مردوخی، فرهاد؛ و شیخ‌احمدی، امیر. (۱۴۰۴). **توسعه سامانه توصیه‌گر هوشمند با رویکرد پالایش گروهی بر پایه مدل‌سازی داده‌های جغرافیایی**، فصلنامه جغرافیا (برنامه‌ریزی منطقه‌ای)، ۱۵(۶۰). ۴۶۴-۴۷۹

DOI:10.22034/jgeoq.2025.551855.4347

* **نویسنده مسئول:** فرهاد مردوخی: **پست الکترونیک:**mardukhi@razi.ac.ir

مقدمه

به صورت کلی، سیستم‌های توصیه‌گر مستقیماً به کاربران کمک می‌کنند تا محتوا، کالاها یا خدمات (مانند کتاب، محصولات دیجیتال، فیلم، موسیقی، برنامه‌های تلویزیونی و...) را با استفاده از جمع‌آوری و تحلیل پیشنهادهای کاربران دیگر بیابند Kim et al., 2010). این سیستم‌ها از آغاز نوشتن اولین مقاله در زمینه سیستم‌های توصیه‌گر پالایش گروهی در اوایل دهه ۹۰ به رشته تحقیقاتی مهمی تبدیل شده‌اند، اما تحقیق بیشتری در زمینه سیستم‌های توصیه‌گر احساس نیاز است؛ چراکه این حوزه هنوز گسترده‌تر و خام‌تر از حوزه‌های تحقیقاتی دیگر است (Park et al.,2012). روش‌های مختلفی برای توصیه به کاربران در این سیستم‌ها به کار گرفته شده است. به دلیل سادگی و نتایج خوب، پالایش گروهی یکی از موفق‌ترین و پراستفاده‌ترین روش‌ها در سیستم‌های توصیه‌گر بوده است(Schafer et al.,2007). تنکی ماتریس امتیازات کاربران به محصولات در پالایش گروهی باعث مشکل شروع سرد می‌شود(Su & Khoshgoftaar,2009).).

از این رو، در یک دهه گذشته اغلب محققان با رویکردهای جدی‌تری به دنبال حل مسائل موجود در سیستم‌های سنتی و به‌کارگیری آن‌ها در سیستم‌های واقعی بوده‌اند؛ به‌خصوص به‌کارگیری روش‌های داده‌کاوی در سیستم‌های توصیه‌گر، در ارائه اطلاعات سفارشی و فردی به کاربر از طریق تحلیل ترجیحاتش مؤثر بوده است (Park et al.,2012). محققان زیادی برای بهبود عملکرد سیستم‌های توصیه‌گر سنتی، برچسب را به این سیستم اضافه کرده‌اند. این برچسب‌ها به‌عنوان پلی بین کاربران و محصولات هستند که به حل مشکل شروع سرد کمک می‌کنند (Zhang et al.,2010)، اما برچسب‌های اضافه‌شده توسط کاربران معمولاً مشکلاتی از قبیل تنک بودن، افزونگی و ابهام را دارند(Sohrabi et al.,2016). برای حل این مشکلات باید به نحوی ساختار زیرین ارتباط بین برچسب‌ها، کشف و همچنین برچسب‌های تکراری کنار گذاشته شوند Sohrabi et al.,2016). همچنین از شبکه‌های عصبی عمیق برای توصیه موسیقی و صفحات وب به کاربران استفاده شده است Sohrabi et al., 2016).

در همین راستا، سوال این تحقیق عبارت است از اینکه آیا عملکرد سیستم توصیه‌گر با استفاده از شبکه عصبی عمیق براساس مدل‌سازیGIS در ایجاد توصیه‌ها از سیستم پالایش گروهی بهتر عمل می‌کند؟ برای پاسخ به این سوال آزمایشی طراحی شده است که در آن، سیستم توصیه‌گر با سیستم توصیه‌گر پالایش گروهی بوسیله مدل‌سازیGIS مقایسه می‌شود.

مدل‌سازی GIS به این فرآیند کمک می‌کند تا الگوهای مکانی و جغرافیایی رفتار کاربران و انطباق آن با ترجیحات محصولات شناسایی شود. این تکنیک می‌تواند داده‌های مرتبط با موقعیت جغرافیایی، نظیر نواحی مورد علاقه کاربران، عادات خرید و الگوهای رفتاری را تحلیل کرده و به سامانه توصیه‌گر کمک کند که پیشنهادات بهتری به کاربران ارائه دهد. این رویکرد می‌تواند به ویژه در یافتن محتواهای مرتبط با خصوصیات جغرافیایی خاص یا نیازهای محلی مؤثر باشد. به این ترتیب، در ادامه مبانی نظری اثر مدل‌سازی GIS بر روی کارایی سیستم‌های توصیه‌گر بررسی خواهد شد و سپس روش تحقیق استفاده شده مرور می‌شود. همچنین تحلیل و بررسی یافته‌ها، و بحث و بررسی نتایج ارائه شده و در انتها، نتیجه‌گیری ارائه خواهد شد.

پیشینه و ادبیات نظری پژوهش

در این بخش ادبیات نظری سه حوزه سیستم‌های توصیه‌گر، یادگیری بازنمایی و الگوریتم LBFGS را بررسی می‌کنیم.

سیستم‌های توصیه‌گر

سیستم‌های توصیه‌گر از توانایی پالایش اطلاعات ضمنی برای پیش‌بینی ترجیح یا عدم ترجیح یک محصول خاص توسط کاربر در زمان تصمیم‌گیری برای انتخاب برخوردار هستند (Motaharinejad et al., 2016). در سال‌های اخیر، نسخه‌های متعددی از سیستم‌های پیشنهاددهنده برای پیشنهاد مفاهیم مختلف در سیستم‌های تجارت الکترونیک و شبکه‌های اجتماعی ارائه شده است(Nez-Valdéz et al., 2012). این سیستم‌ها درباره ترجیحات کاربران در مورد مجموعه‌ای از محصولات (مثل فیلم، موسیقی، کتاب، لطیفه، گیفت‌ها، اپلیکیشن‌ها، وب‌سایت‌ها، مقاصد سفر و محتوای یادگیری آنلاین) اطلاعات جمع‌آوری و از آن‌ها برای توصیه به کاربران استفاده می‌کنند (Bobadilla et al.,2013). سیستم توصیه‌گر می‌تواند از عوامل

جمعیت‌شناختی کاربران (مثل سن، جنسیت و ملیت) استفاده کند. اطلاعات اجتماعی مانند افراد دنبال‌شده، دنبال‌کنندگان، توئیت‌ها، پست‌ها و برچسب‌ها عموماً در سیستم‌های توصیه‌گر وب۲ استفاده شده‌اند(Park et al., 2012) .

در حوزه سیستم‌های توصیه‌گر اجتماعی، کاربران می‌توانند برچسب‌هایی به محصولات اختصاص دهند. یکی از پرکاربردترین روش‌های توصیه به کاربران، روش‌های پالایش گروهی (CF) هستند؛ هرچند که برای بهبود عملکرد، آن‌ها اغلب با دیگر روش‌های پالایش مثل محتوامحور، دانش‌محور و اجتماعی ترکیب می‌شوند (Sarwar et al., 2000) اما این سیستم‌ها معمولاً با مشکل شروع سرد درگیر هستند که زمانی اتفاق می‌افتد که پیش‌نهادهای قابل‌اعتماد به دلیل عدم وجود امتیازدهی کاربران ممکن نباشد (Schafer et al., 2007). مشکل شروع سرد اغلب با استفاده از پالایش ترکیبی بهبود می‌یابد (معمولاً پالایش ترکیبی محتوامحور، پالایش ترکیبی جمعیت‌شناختی و پالایش ترکیبی شبکه اجتماعی (Bengio et al., 2013) برای کاستن مشکل میزان بالای تنکی در پایگاه‌های داده سیستم توصیه‌گر، بعضی مطالعات از روش‌های کاهش بعد، استخراج عوامل و یادگیری بازنمایی استفاده کرده‌اند(Bengio et al., 2009).

یادگیری بازنمایی

یادگیری بازنمایی، یعنی یادگیری بازنمایی‌هایی از داده‌ها که در بیرون کشیدن اطلاعات مفید در هنگام ساختن اپلیکیشن‌های گروه‌بندی یا دیگر پیش‌بینی‌کننده‌ها، بتوانند استفاده شوند(Liu & Nocedal, 1989). روش‌های یادگیری عمیق یکی از انواع راه‌های یادگیری بازنمایی هستند. این روش‌ها از طریق قرار دادن چندین تبدیل‌گر غیرخطی بعد از یکدیگر به دنبال دستیابی به بازنمایی‌های انتزاعی‌تر و درنهایت، مفیدتر هستند.

الگوریتم بهینه‌سازی

L-BFGS یک الگوریتم بهینه‌سازی از خانواده روش‌های شبه‌نیوتونی است که در آن برای جستجو در فضای متغیرها به‌جای استفاده از یک تقریب $n \times n$ (n بعد فضای متغیرها است) از وارون ماتریس هسیان، از یک تقریب که حافظه کمتری اشغال می‌کند، استفاده می‌شود. این الگوریتم تنها چند بردار را ذخیره می‌کند که می‌توانند فضای تقریب زده‌شده را بازنمایی کنند. از این رو، این روش بهینه‌سازی برای مسائل با تعداد متغیر زیاد مناسب است(Marinho & Schmidt-Thieme, 2008).

ادبیات نظری

برای گسترش پالایش گروهی به سامانه‌های برچسب‌زنی اجتماعی، رابطه سه‌تایی بین کاربران، محصولات و برچسب‌ها باید به یک فضا با ابعاد کمتر کاسته شود (Ricci et al., 2011) که اولین مدل از سامانه‌های توصیه‌گر برچسب‌آگاه است. بر اساس این مدل، یک روش توصیه برچسب‌آگاه بر پایه پالایش گروهی در مقاله، Nejdl & Paiu ,2007)ارائه شده است که ابتدا همسایگی کاربر بر اساس تصویری کردن ماتریس کاربر–برچسب محاسبه می‌شود و بعد توصیه بر اساس تجمیع محصولات همسایه‌ای انجام می‌شود. ایده مشابهی در مقاله، Bell & Volinsky (2009) ارائه شده است که ماتریس تصویر کاربر–برچسب به‌منظور محاسبه فهرست رتبه‌بندی‌شده از برچسب‌ها و سپس استخراج توصیه از آن برچسب‌ها به‌دست آمده است(Liang et al,2010). رویکردی ارائه کردند که ماتریس معمول کاربر–محصول را با استفاده از برچسب‌ها تحت عنوان شبه‌کاربران و شبه‌محصولات گسترش‌دهند. سپس یک الگوریتم ترکیبی از ترکیب پالایش گروهی کاربرمحور و پالایش گروهی محصول‌محور ارائه دادند. به‌منظور استفاده بهتر از اطلاعات برچسب‌ها، خوشه‌بندی برچسب‌ها در پالایش گروهی به‌منظور بهبود عملکرد توصیه استفاده شده است. برای رفع مشکل نویز بودن برچسب‌ها، از وزن‌های مختلف برای برچسب‌های مختلف استفاده کرده و آن‌ها را با پالایش گروهی حافظه‌محور سنتی ترکیب کرده‌اند. با بازنمایی رابطه کاربر–محصول–برچسب به‌صورت یک تانسر، ما می‌توانیم ساختار معنایی پسینی شکل داده شده توسط همبستگی‌های سه‌تایی بین آن‌ها را استخراج کنیم. می‌توان این کار را با الگوریتم‌های توصیه بر مبنای تجزیه تانسر انجام داد (Lathauwer et al,2000: 21). این کار در مقاله (Hothoet al,2006: 412) به‌منظور تجزیه تانسر کاربر–محصول–برچسب به ضرب سه ماتریس رتبه پایین و یک تانسر رتبه پایین انجام

شده است. انجام آن بر روی دو مجموعه داده واقعی بهبودهای معنی‌داری در معیارهای ارزیابی دقت و به یادآوری نشان می‌دهد. هونر و همکاران در الگوریتم‌های توصیه گراف‌محور با الهام از الگوریتم جستجوی وب، معروف به الگوریتم Folk rank را ارائه کردند که در آن یک محصول با برچسب‌های مهم که توسط کاربران مهم برچسب زده شده است، خودش مهم می‌شود و این اصل به‌صورت متقارن برای کاربران و برچسب‌ها برقرار است. بر اساس این اصل کاربران، محصولات و برچسب‌ها به‌صورت دوطرفه وزن‌های یکدیگر را تقویت خواهند کرد(Hevneral,2006: 136-215). در نهایت، چندین محصول یا برچسب با وزن‌های بزرگ‌تر توصیه خواهند شد. زایو و دیگران یک سامانه توصیه‌گر برچسبی با استفاده از شبکه‌های عصبی عمیق برای توصیه وب‌سایت‌ها و موسیقی ارائه کرده‌اند. در انتهای پژوهش آن‌ها، به‌کارگیری این الگوریتم برای سایر محصولات، ازجمله فیلم‌های سینمایی، به‌عنوان پیشنهاد برای تحقیقات آتی ارائه شده است (Sohrabi et al,2016).

مدل‌سازی GIS

سیاری از مدل‌سازی دینامیک مکانی–زمانی و رویکردهای تحلیلی که در بخش قبل مورد بحث قرار گرفت، می‌توانند بدون GIS انجام شوند. برای مثال، کرسی و وینکل (۲۰۱۱) روش‌های اکتشافی مختلفی را برای داده‌های مکانی–زمانی، مدل‌های آماری مکانی–زمانی و مدل‌های مکانی–زمانی پویا از دیدگاه‌های آماری ارائه می‌کنند. همچنین انجام شبیه‌سازی مبتنی بر عامل و تجزیه و تحلیل شبکه برای مطالعه دینامیک مکانی–زمانی خارج از یک محیط GIS امکان پذیر است. گودچایلد (۲۰۱۳) چشم اندازهای یک GIS فضا–زمان (STGIS) را بر اساس هفت شکل متمایز از داده‌های فضا–زمان مورد بحث قرار می‌دهد (یعنی ردیابی، توالی‌های زمانی عکس‌های فوری، توالی‌های زمانی پوشش‌های چند ضلعی، اتوماتای سلولی، مدل‌های مبتنی بر عامل، رویدادها و معاملات و داده های چند بعدی). او نتیجه می‌گیرد که، به دلیل تفاوت زیاد در انواع مختلف داده‌های فضا–زمان و سؤالات علمی در هر مورد، بعید است که شاهد ظهور یک GIS فضا–زمان واحد باشیم و «تعدادی از اشکال متمایز STGIS احتمالاً تکامل می‌یابند. بر اساس انواع داده‌های متمایز و مجموعه سوالات علمی (Goodchild, 2013:1076). دو رویکرد کلی برای جفت کردن مدل‌های دینامیکی مکانی–زمانی با GIS وجود دارد. اتصال شل بر اساس تبادل داده بین نرم‌افزار GIS و نرم‌افزار مدل‌سازی خارجی است در حالی که هیچ اشتراکی از پایگاه داده یا توابع مدل‌سازی وجود ندارد. از سوی دیگر، اتصال تنگ، یک محیط توسعه را برای کاربران فراهم می‌کند تا توابع مدل‌سازی پویا مکانی–زمانی را در یک GIS یا اجرای توابع GIS در یک محیط مدل‌سازی پیاده‌سازی کنند (Huang, 2010: 1074).

روش‌شناسی پژوهش

روش تحقیق استفاده شده در این پژوهش از نوع طراحی–آزمایش است. با توجه به مسئله و هدف این پژوهش و نیاز به ایجاد و پیاده‌سازی سامانه توصیه‌گر برچسب‌محور، مراحل فرآیند روش تحقیق علم طراحی به‌عنوان گام‌های اصلی فرآیند تحقیق مورد استفاده قرار گرفته‌اند. این روش تحقیق شامل مراحل شناسایی انگیزه‌ها و مسئله، تعریف اهداف راه‌حل، طراحی و ایجاد، نتیجه‌گیری، ارزیابی و ارتباط است. در مرحله شناسایی، مسئله و انگیزه‌ها باید مسئله تحقیق تعریف و ارزش راه‌حل آن تصدیق شود. در مرحله تعریف اهداف راه‌حل، باید از تعریف مسئله و دانش اینکه چه چیز ممکن و شدنی است، اهداف یک راه‌حل را استخراج کنیم. در مراحل طراحی و ایجاد، محقق مصنوع را خلق می‌کند. این مصنوع‌ها به‌صورت بالقوه ساختارها، مدل‌ها، روش‌ها، نمونه‌ها (Marinho et al., 2011) خصوصیات جدید فنی و اجتماعی یا منابع اطلاعاتی هستند. در مرحله نتیجه‌گیری، کاربرد مصنوع‌در حل یک یا چند نمونه از مسائل نشان داده می شود.

در مرحله ارزیابی، از روش تحقیق آزمایش برای یافتن پاسخ سؤالات تحقیق استفاده شده است و در مرحله ارتباط مسئله و اهمیت آن، مصنوع، کارکرد و بداعت آن، قدرت طراحی آن و تأثیر آن بر محقق و دیگر مخاطبان مرتبط مانند افراد حرفه‌ای نشر می‌یابد. نمودار مراحل تحقیق در شکل ۱ آمده است. تمامی آزمایش‌ها در MATLAB بر روی یک رایانه با پردازنده Intel (R) Core i7 2.00 GHz و ۱۶ گیگابایت حافظه داخلی انجام شده‌اند. در الگوریتم ارائه‌شده، تابع

خط ای شبکه عصبی خودرمزنگار تنک[1]، توسط روش L-EFGS بهینه شده است و برای آن از تابع minFunc که توسط مارک اشمیت[2] ایجاد شده است، استفاده شده است.

پیاده‌سازی سامانه توصیه‌گر

سامانه توصیه‌گر برچسب‌محور با استفاده از شبکه عصبی عمیق در سه مرحله مختلف طراحی و ایجاد شده است. خلاصه مراحل در شکل زیر آورده شده‌اند:

شکل (۱): نمودار مراحل پیاده‌سازی

مرحله ۱: پیکربندی داده‌ها

مرحله ۱، داده‌ها را در یک پیکربندی مناسب جهت استفاده در مراحل بعدی آماده می‌کند و آن‌ها را به دو قسمت داده‌های تعلیم (۸۰٪ داده‌ها) و داده‌های آزمون (۲۰٪ داده‌ها) تقسیم می‌کند. پیکربندی مناسب شامل دو جدول کاربر –برچسب و کاربر–محصول است. سطر اول این جدول همه شناسه‌های برچسب‌های موجود و ستون اول آن آیدی‌های کاربران است. اگر کاربری فیلمی را برچسب زده باشد، در سلولی از این جدول که از تقاطع سطر آیدی کاربر و ستون آیدی برچسب به‌دست‌آمده است، عدد یک و در غیر این صورت، عدد صفر قرار داده می‌شود. برای کمتر شدن میزان و زمان محاسبات، تنها برچسب‌هایی که درمجموع همه برچسب‌ها بیشتر از چهار بار تکرار شده‌اند، در این جدول نگه داشته شده‌اند و برچسب‌های با کمتر ازسه بار تکرار از جمع برچسب‌ها حذف شده‌اند. جدول کاربر–محصول نیز تقریباً به همین صورت ساخته می‌شود.

مرحله ۲: تعلیم سامانه، تحلیل عاملی و استخراج عوامل

در این بخش ما قصد داریم با تحلیل عاملی برچسب‌ها و استخراج عوامل، بعد فضایی را که سامانه توصیه‌گر بر روی آن عمل می‌کند، کاهش دهیم. پروفایل هر کاربر u به‌صورت ستون دوم الی ستون آخر سط ر u اُم ماتریس کاربر–برچسب تعریف (پروفایل شامل آیدی کاربر نیست) و با ( ) نمایش داده می‌شود. اگر m کاربر آزمون و n برچسب داشته باشیم، پروفایل هریک از آن‌ها به‌صورت زیر نشان داده می‌شود:

---

[1]. sparse auto-encoder

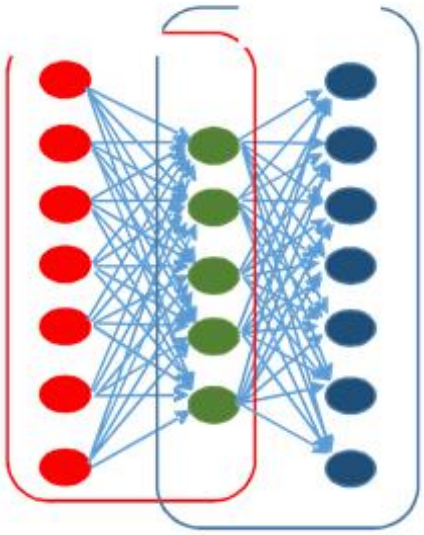
[2]. https:// sccn. ucsd. edu/ svn/ software/ tags/ EEGLAB7_0_2_5 beta/ external/ fieldtrip 20090727/ classification/ toolboxes/ external/minFunc/minFunc.m

$$( ) = \quad ( ) , ( ) , \dots , ( ) \quad = 1, 2, 3, \dots , \qquad (1)$$

$$( ) = 1 \text{ یا } 0 \text{ و } \quad = 1, 2, \dots , \quad \text{و} \quad = 1, 2, \dots , \qquad (2)$$

برای تحلیل عاملی از خودرمزنگاری تنک استفاده می‌کنیم که یک شبکه عصبی سه لایه‌ای شامل لایه ورودی، یک لایه پنهان و یک لایه خروجی بوده و بین همه گره‌های یک لایه با گره‌های لایه بعدی یال وجود دارد. بعد لایه ورودی با تعداد برچسب‌ها برابر است و پروفایل هریک از کاربران به‌عنوان ورودی به این لایه داده می‌شود. بعد لایه پنهان برابر تعداد عواملی است که قصد داریم آن‌ها را استخراج کنیم. بعد لایه خروجی با بعد لایه ورودی یکسان و برابر تعداد برچسب‌ها است. معماری این شبکه در شکل (۲) نشان داده شده است.

رمزگشا-رمزنگار



شکل (۲): معماری شبکه عصبی با یک لایه پنهان

در لایه خروجی، ما قصد داریم بردار ورودی را دوباره بازتولید کنیم. هرچقدر میزان اختلاف بردار ورودی و بردار خروجی‌کمتر باشد، عوامل استخراج شده عوامل بهتری هستند. خودرمزنگار تنک شامل دو بخش رمزنگار و رمزگشا است. بخش رمزگشا شامل لایه ورودی و لایه پنهان است. این لایه با ماتریس وزن با لایه پنهان ارتباط دارد. خروجی لایه پنهان برابر فعال‌سازی گره‌های لایه پنهان است؛

بنابراین، خروجی لایه پنهان برای کاربر u برابر است با:

$$( ) = (\quad \times ( ) +) \qquad (3)$$

که در آن ( )خروجی گره i اُم لایه پنهان برای ورودی پروفایل کاربر u، ( ) پروفایل کاربر u و بردار b بایاس لایه ورودی است. K تعداد گره‌های لایه پنهان، n تعداد فیلم‌ها وتابع سیگموئید است. بخش رمزگشا شامل لایه پنهان و لایه خروجی است که تلاش می‌کند ورودی به لایه اول شبکه (پروفایل کاربران) را بازسازی کند. خروجی لایه خروجی که خروجی لایه رمزگشا هم هست برابر است با:

$$( ) = (\quad \times ( ) +) \qquad (4)$$

که در آن ( ) خروجی شبکه عصبی و خروجی رمزگشا برای کاربر u، ماتریس وزن بین لایه پنهان و لایه خروجی و c بایاس لایه دوم است.

( ) بردار بازنمایی پروفایل ( )کاربر u است و بر اساس عوامل استخراج شده که خروجی‌های لایه پنهان شبکه عصبی هستند، نوشته شده است.

ما سعی می‌کنیم در تعلیم شبکه، تفاوت بین ورودی به شبکه و خروجی شبکه کمترین مقدار ممکن شود. برای این کار یک تابع خطا تعریف می‌کنیم و آن را روی مولفه‌های این تابع که ماتریس وزن است، کمینه می‌کنیم.

به خاطر تقارن شبکه مقادیر برابرکه برابر با ترانهاده ماتریس است، انتخاب می‌شود. با اضافه کردن واگرایی‌کولبک–لیبلر[1] به خطا به خاطر تنکی داده‌های ورودی، مقدار خطا برابر است با:

$$1 \quad ( ) - \underline{( )} \quad +$$

$$2$$

$$— \qquad (5)$$

$$1$$

$$( ) - (\quad \times \quad ( ) + \quad +) \quad + \qquad .۱$$

$$2$$

که در آن:

$$(\quad) \qquad (6) \qquad =$$

$$1$$

$$(\mid \quad = \qquad + (1 -) \quad 1 \qquad (7)$$

و

$$1 \quad ( ); \quad , \qquad (8) =$$

که در آن mبرابر است با تعداد داده‌های تعلیم. در رابطه بالا، مقادیر ( )ها همان ورودی‌های داده‌های تعلیم و مشخص هستند و متغیرهایی که باید تابع فوق ر ا روی آن‌ها کمینه کنیم، درایه‌های ماتریس‌و بردارهای b و c هستند.

بنابراین داریم:

$$\min (\quad . .) \qquad — \qquad (9)$$

برای کمینه‌کردن خطا، از الگوریتم l-bfgs پیاده‌سازی شده در تابع minFunc استفاده کردیم.

از این بهینه‌سازی ماتریس وزن‌های بهینه و b برای ساختن بهترین عوامل از پروفایل کاربران به دست می‌آید. با استفاده از این وز ن‌های بهینه، یک نگاشت

$\rightarrow$ =: از فضای برچسب‌ها به فضای عوامل استخراج شده، به‌صورت زیر تعریف می‌کنیم:

$$:= (\quad +) \qquad (10)$$

با این نگاشت تمام پروفایل‌های جدول‌های آزمون و تعلیم را به‌روزرسانی می‌کنیم تا از آن‌ها در مرحله ۳ برای تعیین میزان شباهت کاربران با کاربر هدف و توصیه به او استفاده کنیم.

مرحله ۳: توصیه به کاربر

در این قسمت، قصد داریم با استفاده از کاربران تعلیم به یک کاربر u که کاربر هدف نامیده می‌شود و از کاربران آزمون انتخاب شده، توصیه کنیم. در این مرحله ابتدا میزان شباهت پروفایل تصویر شده در فضای عوامل استخراج شده‌کاربر هدف با پروفایل تصویرشده دیگر کاربران مقایسه می‌شود. کاربران‌برحسب این میزان شباهت مرتب می‌شوند و تعدادی برابر با متغیر همسایگی که از قبل تعیین شده از بین آن‌ها انتخاب می‌شود. از ضرب میزان شباهت هریک از کاربران همسایگی در بردار کاربر–فیلم آن‌ها و جمع آن‌ها با یکدیگر میزان امتیاز پیش‌بینی‌شده هریک از فیلم‌ها برای کاربر هدف تعیین می‌شود و از بین آن‌ها به تعداد طول فهرست توصیه فیلم‌های با بیشترین امتیاز انتخاب می‌شوند.

برای تعیین میزان شباهت بین کاربر هدف u که از بین کاربران آزمون انتخاب شده است و سایر کاربران تعلیم از معیار شباهت کسینوسی استفاده می‌کنیم. میزان شباهت بین کاربر u و کاربر v از رابطه زیر به دست می‌آید.

$$< \quad . \quad > \qquad (11) =$$

$$\| \quad \| \times \| \quad \| \quad -۱ \qquad _____$$

که در آن و پروفایل‌های تصویر شده در فضای عوامل هستند.

1-Kullback-Liebler

< > نشان‌دهنده مجموع ضرب تک‌به‌تک درایه‌های متناظر دو بردار و || || نشان‌دهنده اندازه یک بردار هستند. به مجموعه کاربرانی که قرار است در توصیه به کاربر u شرکت کنند همسایگی این کاربر گفته می‌شود و تعداد آن‌ها را با متغیر مشخص می‌کنیم. میزان امتیاز پیش‌بینی‌شده هریک از فیلم‌ها برای کاربر u به‌صورت زیر مشخص می‌شود.

× ( ) (12) =

. . .

که در آن، . ( ) سلول سطر مربوط به کاربر v و ستون فیلم با آیدی i جدول کاربر–فیلم است. به‌عبارت دیگر، . ( ) برابر یک است، اگر کاربر v فیلم با آیدی i را برچسب‌گذاری کرده باشد و در غیر این صورت، صفر است.

خروجی فرآیند فوق یک بردار است که تعداد درایه‌های آن به تعداد فیلم‌ها است، اگر سطر اول جدول کاربر–فیلم را به بالای این بردار اضافه کنیم جدولی تشکیل می‌شود که به ما نشان می‌دهد امتیاز پیش‌بینی‌شده هر فیلم برای کاربر u چقدر است. بسته به مقدار طول فهرست توصیه فیلم‌های با بیشترین امتیاز برای توصیه به کاربر انتخاب می‌شوند.

آزمایش

این موضوع هستند که از بین کاربر است و چه تعداد از کل این معیارهای ارزیابی سامانه معیارهای دقت و به‌یادآوری به ترتیب نشان‌دهنده فیلم‌های توصیه شده چه تعداد از آن‌ها موردعلاقه فیلم‌های موردعلاقه کاربر به او توصیه شده است که توسط روابط زیر محاسبه می‌شوند.

۱- دقت = تعداد کل فیلم‌های توصیه شده به کاربر / فیلم‌های توصیه شده که کاربر به آن‌ها امتیاز مثبت داده است

۱. به‌یادآوری = تعداد کل فیلم‌هایی که کاربر به آن‌ها امتیاز مثبت داده است / فیلم‌های توصیه شده که کاربر به آن‌ها امتیاز مثبت داده است

۲.

معیار دقت f1 در ترکیبی از دو معیار فوق است و از رابطه ۱۶ به‌دست می‌آید.

۱- یادآوری + دقت = ۱ ۲ × یادآوری × دقت

در آزمایش طراحی‌شده مقادیر معیارهای ارزیابی دقت، به‌یادآوری و f1 سامانه برچسب‌محور بر مبنای شبکه عصبی عمیق با یک لایه پنهان با ۱۰۰۰ گره برای داده‌های وب‌سایت www.movielens.org با سامانه پالایش گروهی برای طو ل‌های مختلف فهرست توصیه با یکدیگر مقایسه شده‌اند. مجموعه داده این سامانه شامل ۱۷۱۵ کاربر، ۵۴۶۰ فیلم و ۱۹۳۴برچسب است که هر برچسب حداقل چهار بار تکرار شده است.

مولفه‌های زیر برای هر دو سامانه یکسان در نظر گرفته شده است.

α. : 0.1فعال‌سازی متوسط

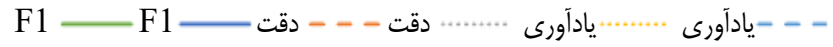
۲.: 3وزن جمله جریمه تنکی

ι. : 90سایز همسایگی

یافته‌ها

در شکل زیر نمودارهای دقت، به یادآوری و f1برای دو سامانه برچسب‌محور بر مبنای شبکه عصبی عمیق (مدل برچسبی) و سامانه پالایش گروهی آورده شده است.

طول فهرست توصیه



شکل (۳): مقایسه سامانه برچسب‌محور و پالایش گروهی

همین‌طور که در شکل ۳ مشخص است، سامانه برچسب‌محور بر مبنای شبکه عصبی عمیق در هر سه معیار ارزیابی دقت، به یادآوری و F1 برای تمام طول فهرست‌های توصیه از یک تا ۲۰۰ مقدار بیشتری را در توصیه فیلم‌های سینمایی به نسبت سامانه پالایش گروهی سنتی کسب کرده است. از شکل بالا مشخص است که نمودار یادآوری، برای هر دو سامانه صعودی است. علت این موضوع آن است که با افزایش طول فهرست توصیه، تعداد فیلم‌هایی از کاربر که قبلاً مورد علاقه او بوده‌اند و حالا هم توصیه شده‌اند، بیشتر می‌شود؛ یعنی ازنقطه‌نظر معیار یادآوری، هرچقدر تعداد فیلم‌های توصیه‌شده بیشتر باشد، سامانه بهتر عمل می‌کند. نمودار دقت برحسب طول فهرست توصیه برای هر دو سامانه نزولی است؛ چراکه هرچقدر طول فهرست توصیه بیشتر می‌شود، تعداد فیلم‌های غلط توصیه‌شده هم بیشتر می‌شوند که این موضوع به کاهش دقت منجر می‌شود. با افزایش طول فهرست توصیه، دقت دو سامانه‌به یکدیگر نزدیک می‌شوند؛ چراکه سامانه توصیه‌گر برچسبی به خاطر زیادشدن تعداد توصیه‌ها برتری خود را در دقت از دست می‌دهد. معیار ارزیابی f1 که ترکیبی از دو معیار دقت و یادآوری است، یک حداکثر دارد؛ بنابراین طول فهرست توصیه نزدیک به ۱۵، بهترین عملکرد برای سامانه توصیه‌گر برچسب‌محور را داراست.

بحث و بررسی

حرکت از GIS لحظه‌ای استاتیک به GIS زمانی، که توانایی مدیریت انواع دینامیک‌های مکانی–زمانی را بهبود می‌بخشد، توجه قابل‌توجهی را در تحقیقات جلب کرده است. مدل‌های داده متعددی برای GIS زمانی پیشنهاد شده‌اند، اما توسعه یک چارچوب یکپارچه علم GIS برای ادغام رویکردها و تعاریف مختلف فضا–زمان همچنان یک چالش تحقیقاتی است. مقایسه تفاوت‌های دینامیک مکانی–زمانی بین مکان‌های مختلف نیز مفید است. این پویایی‌ها در مقیاس‌های مختلف (مانند فردی، محلی، میانی و جهانی) و از ابعاد چندگانه (مانند الگوی مکانی، روند زمانی و توزیع آماری) قابل‌بررسی است. ترکیب سیستماتیک این ابعاد و مقیاس‌ها می‌تواند چارچوب تحلیلی جامعی برای شناسایی شکاف‌ها و مرزهای تحقیقاتی ایجاد کند (Ye and Rey, 2013). استفاده از داده‌های باز و جعبه ابزارهای منبع‌باز با زبان‌های برنامه‌نویسی مانند R و Python، تغییر پارادایمی به‌سوی همکاری بین‌رشته‌ای قوی در تحقیقات دینامیک مکانی–زمانی ایجاد کرده است.

در این راستا، این پژوهش به بررسی عملکرد سامانه توصیه‌گر برچسب‌محور مبتنی بر مدل‌سازی GIS برای توصیه فیلم‌های سینمایی پرداخته است. پرسش اصلی تحقیق این است که آیا سامانه توصیه‌گر مبتنی بر شبکه‌های عصبی عمیق و مدل‌سازی GIS نسبت به سامانه پالایش گروهی عملکرد بهتری دارد؟ برای پاسخ به این پرسش، آزمایشی طراحی شد که در آن سامانه توصیه‌گر برچسب‌محور با سامانه پالایش گروهی در معیارهای ارزیابی دقت، به‌یادآوری و F1 مقایسه شد. داده‌های این آزمایش از وب‌سایت movielens.org، یکی از سایت‌های معتبر توصیه فیلم، استخراج شد. نتایج نشان داد که سامانه برچسب‌محور در هر سه معیار ارزیابی برای طول‌های مختلف لیست توصیه (از ۱ تا ۲۰۰) عملکرد بهتری نسبت به سامانه پالایش گروهی دارد. دلیل این برتری، توانایی سامانه برچسب‌محور در رفع مشکل تنکی داده‌ها با استفاده از یک خودرمزنگار تنک برای استخراج عوامل انتزاعی‌تر از برچسب‌ها است. این عوامل، شناسایی دقیق‌تر کاربران مشابه با کاربر هدف را امکان‌پذیر کرده و توصیه‌های دقیق‌تری ارائه می‌دهند.

در سامانه‌های توصیه‌گر مدل‌محور، ابتدا مدلی بر اساس داده‌ها ساخته می‌شود و سپس توصیه‌ها بر مبنای آن مدل انجام می‌گیرند. این مدل‌ها به‌صورت برخط به‌روزرسانی نمی‌شوند و هرچند وقت یک‌بار بازسازی می‌شوند، که این امر زمان توصیه را نسبت به پالایش گروهی کاهش می‌دهد؛ زیرا تعداد عوامل استخراج شده کمتر از تعداد برچسب‌های اولیه است. مقایسه با تحقیقات پیشین نشان می‌دهد که سامانه پیشنهادی در معیار به‌یادآوری عملکرد بهتری نسبت به دقت دارد. به‌عنوان مثال، مارینو یک سامانه ترکیبی کاربرمحور و محصول‌محور با برچسب برای توصیه موسیقی طراحی کرد که به‌یادآوری ۰.۳۴ برای لیست‌های ده‌تایی گزارش داد، اما سامانه برچسب‌محور مبتنی بر شبکه عصبی عمیق در این معیار بهتر عمل می‌کند. لیانگ و همکاران نیز سامانه‌ای با برچسب‌های وزنی برای توصیه کتاب و وب‌سایت طراحی کردند که به‌یادآوری کمتری نسبت به سامانه پیشنهادی داشت، اما دقت بیشتری نشان داد. به‌طور کلی، رابطه معکوسی بین به‌یادآوری و دقت وجود دارد؛ سامانه‌هایی با به‌یادآوری بالاتر، تعداد بیشتری از محصولات موردعلاقه کاربر را توصیه می‌کنند، اما ممکن است فیلم‌هایی را نیز پیشنهاد دهند که موردتوجه کاربر نباشند، که این امر دقت را کاهش می‌دهد. در مجموع، این سامانه در مقایسه با تحقیقات پیشین، در به‌یادآوری برتری دارد، اما در دقت نسبت به برخی سامانه‌ها عملکرد ضعیف‌تری نشان می‌دهد.

نتیجه‌گیری و پیشنهادها

این پژوهش به طراحی و ارزیابی سامانه توصیه‌گر برچسب‌محور مبتنی بر شبکه‌های عصبی عمیق و مدل‌سازی GIS برای توصیه فیلم‌های سینمایی پرداخته است، که برای اولین بار از سامانه‌ای که پیش‌تر برای توصیه موسیقی و وب‌سایت‌ها توسعه یافته بود، در حوزه فیلم‌های سینمایی استفاده شد. از منظر مدل‌سازی GIS، حرکت از GIS لحظه‌ای استاتیک به GIS زمانی، که توانایی مدیریت دینامیک‌های مکانی–زمانی را در مقیاس‌های مختلف (فردی، محلی، میانی و جهانی) و ابعاد چندگانه (الگوی مکانی، روند زمانی و توزیع آماری) بهبود می‌بخشد، نقش مهمی در این سامانه ایفا کرده است. ترکیب سیستماتیک این ابعاد و مقیاس‌ها چارچوبی تحلیلی برای شناسایی شکاف‌های تحقیقاتی فراهم می‌کند. استفاده از داده‌های باز و جعبه‌ابزارهای منبع‌باز مانند R و Python نیز همکاری بین‌رشته‌ای قوی‌تری را در توسعه این سامانه تسهیل کرده است. در سامانه‌های مدل‌محور، مدلی بر اساس داده‌ها ساخته شده و توصیه‌ها بر مبنای آن انجام می‌شود. این مدل‌ها به‌صورت برخط به‌روزرسانی نمی‌شوند و هرچند وقت یک‌بار بازسازی می‌شوند، که زمان توصیه را نسبت به پالایش گروهی کاهش می‌دهد، زیرا تعداد عوامل استخراج شده کمتر از برچسب‌های اولیه است.

مقایسه با تحقیقات پیشین نشان می‌دهد که این سامانه در معیار به‌یادآوری عملکرد بهتری نسبت به دقت دارد. برای مثال، مارینو سامانه‌ای ترکیبی برای توصیه موسیقی طراحی کرد که به‌یادآوری ۰.۳۴ برای لیست‌های ده‌تایی گزارش داد، اما سامانه پیشنهادی در این معیار بهتر عمل می‌کند. لیانگ و همکاران نیز سامانه‌ای با برچسب‌های وزنی برای توصیه کتاب و وب‌سایت طراحی کردند که به‌یادآوری کمتری داشت، اما دقت بیشتری نشان داد. رابطه معکوس بین به‌یادآوری و دقت نشان می‌دهد که سامانه‌های با به‌یادآوری بالاتر، محصولات بیشتری را که موردعلاقه کاربر هستند توصیه می‌کنند، اما ممکن است مواردی را نیز پیشنهاد دهند که موردتوجه کاربر نباشند، که این امر دقت را کاهش می‌دهد.

از نظر کاربردی، این سامانه در هر سیستمی که کاربران در آن برچسب‌گذاری می‌کنند، مانند شبکه‌های اجتماعی برای توصیه کاربران بر اساس شباهت برچسب‌ها، فروشگاه‌های آنلاین با امکان برچسب‌گذاری کالا، سیستم‌های مدیریت دانش برای برچسب‌گذاری اسناد و موتورهای جستجو قابل‌استفاده است. این سامانه می‌تواند به‌صورت آزمایشی در کنار سامانه‌های فعلی پیاده‌سازی شود و با رفع اشکالات احتمالی به کار گرفته شود. همچنین، در اپلیکیشن‌های فروش آنلاین، شبکه‌های اجتماعی یا پلتفرم‌های تولید محتوا برای توصیه محتوای مناسب کاربرد دارد.

محدودیت‌های این پژوهش شامل عدم ارزیابی زمان اجرای الگوریتم‌ها به دلیل نبود مسئله‌ای با مقیاس عظیم در دامنه تحقیق است. پیشنهاد می‌شود در تحقیقات آتی، نمونه داده‌های مناسب برای ارزیابی این موضوع استفاده شود. همچنین، بسامد تکرار برچسب‌ها توسط کاربران در این سامانه لحاظ نشده است، که می‌تواند در آینده مدنظر قرار گیرد. علاوه بر این، امتیازهای ۱ تا ۵ کاربران برای فیلم‌ها در توصیه‌ها استفاده نشده است و پیشنهاد می‌شود در تحقیقات بعدی این امتیازها در کنار برچسب‌ها برای تعیین شباهت کاربران به کار روند. آزمایش سامانه با مجموعه داده‌های بزرگ‌تر و استفاده از روش‌های تحلیل بزرگ‌داده مانند

هادوپ یا تقلیل نگاشت نیز می‌تواند عملکرد سامانه را در مقیاس‌های بزرگ‌تر بررسی کند. تحلیل حس‌کاوی برای تعیین بار احساسی برچسب‌ها و استفاده از آن در توصیه‌ها، بررسی عملکرد سامانه برای کالاهای دیگر مانند کتاب یا فروشگاه‌های ترکیبی و ارزیابی واقع‌بینانه‌تر معیارهای سامانه از طریق پرسشنامه نیز از دیگر پیشنهادات برای تحقیقات آتی است. در نهایت، استفاده از الگوریتم‌های ژنتیکی، روش‌های فازی یا شبکه‌های عصبی با معماری متفاوت می‌تواند با نتایج این پژوهش مقایسه شود.

منابع

Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. Knowledge-Based Systems, 46, 109–132.

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1798–1828.

Bengio, Y. (2009). Learning deep architectures for AI. Foundations and Trends in Machine Learning, 2(1), 1–127.

De Lathauwer, L., De Moor, B., & Vandewalle, J. (2000). A multilinear singular value decomposition. SIAM Journal on Matrix Analysis and Applications, 21(4), 1253–1278.

Firan, C. S., Nejdl, W., & Paiu, R. (2007). The benefit of using tag-based profiles. In LA-WEB 2007: Latin American Web Conference (pp. 32–41). IEEE.

Goodchild, M. F. (2013). Geospatial and temporal perspectives of a GIS. Annals of the American Association of Geographers, 103(5), 1072–1077.

Huang, B., Wu, B., & Barry, M. (2010). Geographically and temporally weighted regression for modeling spatial-temporal variation of housing prices. International Journal of Geographical Information Science, 24(3), 383–401.

Hotho, A., Jäschke, R., Schmitz, C., & Stumme, G. (2006). Information retrieval in folksonomies: Search and ranking. In European Semantic Web Conference (pp. 411–426). Springer.

Hevner, A., & Chatterjee, S. (2010). Design research in information systems: Theory and practice (Vol. 22, pp. 136–215). Springer.

Kim, H. N., Ji, A. T., Ha, I., & Jo, G. S. (2010). Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation. Electronic Commerce Research and Applications, 9(1), 73–83.

Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. Computer, 42(8), 30–37.

Liang, H., Xu, Y., Li, Y., Nayak, R., & Tao, X. (2010). Connecting users and items with weighted tags for personalized item recommendations. In Proceedings of the 21st ACM Conference on Hypertext and Hypermedia (pp. 51–60). ACM.

Liu, D. C., & Nocedal, J. (1989). On the limited memory BFGS method for large scale optimization. Mathematical Programming, 45(1–3), 503–528.

Motaharinejad, M., Zolfagharzadeh, M., Khadangi, E., & Sadabadi, A. (2016). Designing a model for improving banking recommender systems based on predicting customers' interests: Application of data mining techniques. Journal of Information Technology Management, 8(2), 393–314.

Marinho, L. B., & Schmidt-Thieme, L. (2008). Collaborative tag recommendations. In Data Analysis, Machine Learning and Applications (pp. 533–540). Springer.

Núñez-Valdéz, E. R., Lovelle, J. M. C., Martínez, O. S., García-Díaz, V., de Pablos, P. O., & Marín, C. E. M. (2012). Implicit feedback techniques on recommender systems applied to electronic books. Computers in Human Behavior, 28(4), 1186–1193.

Park, D. H., Kim, H. K., Choi, I. Y., & Kim, J. K. (2012). A literature review and classification of recommender systems research. Expert Systems with Applications, 39(11), 10059–10072.

Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to recommender systems handbook. In Recommender Systems Handbook (pp. 5–35). Springer.

Schafer, J. B., Frankowski, D., Herlocker, J., & Sen, S. (2007). Collaborative filtering recommender systems. In The Adaptive Web (pp. 291–324). Springer.

Su, X., & Khoshgoftaar, T. M. (2009). Survey of collaborative filtering techniques. Advances in Artificial Intelligence, 2009, Article 421425.

Sohrabi, B., Raeesi Vanani, I., & Zareh Mirkabad, F. (2016). Designing a recommender system for optimizing and managing bank facilities through the utilization of clustering and classification algorithms. Modern Research in Decision Making, 1(2), 53–76.

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2000). Application of dimensionality reduction in recommender system: A case study (Technical Report No. TR-00-043). University of Minnesota, Department of Computer Science.

Ye, X., & Rey, S. J. (2013). A framework for exploratory space–time data analysis of economic data. The Annals of Regional Science, 50(1), 315–339.

Zhang, Z. K., Liu, C., Zhang, Y. C., & Zhou, T. (2010). Solving the cold-start problem in recommender systems with social tags. EPL (Europhysics Letters), 92(2), 28002.

Zuo, Y., Zeng, J., Gong, M., & Jiao, L. (2016). Tag-aware recommender systems based on deep neural networks. Neurocomputing, 204, 51–60.