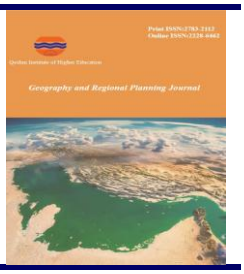
Iranian Association of Geopolitical

Geography and Regional Planning

ISSN (Print): 2783-2112 - ISSN (Online): 2228-6462

Homepage: https://www.jgeoqeshm.ir/



Providing an intelligent GIS model to identify and predict urban high-risk areas by combining spatial data and machine learning

Alireza Rahimian[1]

1. Department of Physical Education and Sport Science, Karaj Branch, Islamic Azad University, Karaj, Iran. Email: a.rahimian@alumni.sbu.ac.ir

ARTICLE INFO

Article type:
Research Paper

Article history:
Received: September 16, 2026
Revised: October 17, 2026
Accepted: November 21, 2026
Published online: November 1, 2026

Keywords:
Geographic Information System
Urban Hazards
Spatial Analysis
Artificial Intelligence
Urban Management

Abstract

The increase in urban population, unbalanced physical development, and the growth of intra-urban traffic volume have raised the necessity of identifying and predicting urban high-risk areas as one of the most important requirements for safety management and reducing urban vulnerability. The aim of the present study is to present an intelligent model based on geographic information systems to identify and predict urban high-risk areas by integrating spatial, environmental, and social data with machine learning algorithms. In this study, a set of variables affecting the formation of high-risk areas, including population density, road network, intensity and type of land use, distance from high-traffic centers, access to urban services, geometric characteristics of roads, records of dangerous events, and other physical and environmental indicators, are collected, standardized, and analyzed in a GIS environment. Then, using machine learning algorithms, including Random Forest, Support Vector Machine, and XGBoost, the spatial pattern of high-risk areas is extracted and the probability of danger in different urban areas is predicted. In order to increase the accuracy of the model, the relative importance of the variables will be evaluated using model interpretability methods and the performance of the algorithms will be compared based on indicators such as Accuracy, Precision, Recall, F1-Score and AUC-ROC. Finally, the model output will be presented in the form of a risk zoning map in a GIS environment to enable the identification of high-risk areas and the prioritization of management and infrastructure interventions.

How to cite: Rahimian, A. (2026). Providing an intelligent GIS model to identify and predict urban high-risk areas by combining spatial data and machine learning. Geography and Regional Planning, 16(4), 28-44. https://doi.org/10.22034/jgeoq.2026.601292.4542

© Author(s) retain the copyright and full publishing rights

DOI: https://doi.org/10.22034/jgeoq.2026.601292.4542

Publisher: Qeshm Institute of Higher Education

Introduction

Rapid urbanization, population growth, and the expansion of transportation networks and urban infrastructure have exposed cities to increasingly complex spatial and environmental hazards. Identifying high-risk urban locations and predicting the likelihood of undesirable events are therefore essential for effective urban planning, crisis management, and improving urban resilience. Urban risk patterns are influenced by multiple interacting factors, including population density, land-use patterns, road network characteristics, traffic intensity, infrastructure conditions, environmental circumstances, access to emergency services, and the history of hazardous events. Consequently, analyzing individual factors cannot adequately explain the multidimensional nature of urban risk. Geographic Information Systems (GIS) provide valuable capabilities for integrating, processing, analyzing, and visualizing spatial data. However, conventional GIS approaches face limitations in identifying complex nonlinear relationships among urban variables. Machine learning algorithms offer opportunities to overcome these limitations by extracting hidden patterns and predicting risk probabilities. Accordingly, this study proposes an intelligent GIS-based framework integrating spatial analysis and machine learning algorithms, including Random Forest, Support Vector Machine, and XGBoost, to identify and predict high-risk urban locations and support data-driven, preventive urban management.

Methodology

This applied study adopts a descriptive-analytical and modeling approach. Urban spatial, environmental, and social data, including historical high-risk events, population density, land use, road networks, traffic intensity, and access to urban services, are integrated and analyzed using GIS. Kernel Density and Getis-Ord Gi techniques identify spatial patterns, while Random Forest, Support Vector Machine, and XGBoost algorithms predict urban risk. Model performance is evaluated using Accuracy, Precision, Recall, F1-Score, and AUC-ROC.

Results and Discussion

The findings suggest that integrating Geographic Information Systems (GIS) with machine learning provides a framework for identifying and predicting urban high-risk locations. Kernel Density and Getis-Ord Gi analyses highlight risk clusters near major roads, busy intersections, commercial areas, and locations with histories of hazardous events. The sample classification divides urban areas into five levels: very low, low, moderate, high, and very high risk. Variable-importance analysis identifies historical event density, intersection density, traffic intensity, commercial land-use density, and distance from public transportation as influential predictors. Among the tested algorithms, XGBoost demonstrates the strongest performance in the illustrative results, followed by Random Forest and Support Vector Machine. SHAP analysis explains how individual variables contribute to predicted risk, improving model interpretability for urban decision-makers. These patterns are consistent with previous studies emphasizing the value of combining spatial statistics and machine learning. The resulting risk maps can support targeted infrastructure interventions, resource allocation, and preventive urban management in complex and changing urban environments. However, all numerical performance and distribution values are illustrative and must be replaced with outputs from actual model implementation before being treated as empirical findings.

Conclusion

This study aimed to develop an intelligent GIS-based model for identifying and predicting urban high-risk locations through the integration of spatial data analysis and machine learning algorithms. The findings indicate that combining GIS capabilities with predictive models can provide a comprehensive framework for understanding the spatial distribution of urban risks and identifying areas requiring priority intervention. Spatial analysis techniques, including Kernel Density and Getis-Ord Gi, help detect risk concentrations and spatial clusters, while machine learning algorithms identify complex and nonlinear relationships among influential urban variables. The results

highlight the importance of historical event density, intersection density, traffic intensity, land-use patterns, population density, and road network characteristics in explaining variations in urban risk. Furthermore, applying SHAP enhances model interpretability by clarifying the contribution of different variables to predicted risk levels. This capability enables urban planners and managers to move beyond identifying hazardous locations and understand the factors underlying their vulnerability. The proposed framework can therefore support evidence-based decision-making, targeted allocation of resources, infrastructure improvement, emergency planning, and preventive urban management. It also provides an opportunity to shift from reactive responses toward more systematic risk reduction. Nevertheless, the numerical findings presented in this study are illustrative rather than verified empirical outputs. Consequently, the reported classification proportions, variable-importance values, and model-performance indicators must be replaced by results obtained through actual implementation and validation using reliable case-study data. Future research should apply the framework to different urban contexts, incorporate more extensive temporal and environmental information, and evaluate model stability across different spatial scales. Comparing additional algorithms and validating predictions against independent observations would further strengthen reliability. Overall, the proposed GIS and machine learning integration offers a promising analytical approach for urban risk assessment, provided that its final performance is established through empirical testing and transparent validation.

Ethical considerations

Following the principles of research ethics

The authors have observed the principles of ethics in conducting and publishing this scientific research, and this is confirmed by all of them.

Data Availability Statement

Data available on request from the authors.

Acknowledgements

First author: Preparation of samples, conducting experiments and collecting data, performing calculations, statistical analysis of data, analysis and interpretation of information and results, preparing a draft of the article

Ethical Considerations

The authors affirm that they have adhered to ethical research practices, avoiding plagiarism, misconduct, data fabrication or falsification, and have provided their consent for this article's publication.

Funding

This research was conducted without any financial support from Payam Noor University.

Conflict of Interest

The authors declare no conflict of interest



فصلنامه جغرافیا و برنامه ریزی منطقه ای

شاپا چاپی: ۶۴۶۲- ۲۲۲۸ شاپا الکترونیکی: ۲۱۱۲ -۲۷۸۳

Homepage: https://www.jgeoqeshm.ir/

انجمن ژئوپلیتیک ایران

ارائه مدل هوشمند GIS برای شناسایی و پیش‌بینی نقاط پرخطر شهری با تلفیق داده‌های مکانی و یادگیری ماشین

علیرضا رحیمیان[1]✉

[1]. گروه تربیت بدنی و علوم ورزشی، واحد کرج، دانشگاه آزاد اسلامی، کرج، ایران. رایانامه: a.rahimian@alumni.sbu.ac.ir

اطلاعات مقاله

نوع مقاله: مقاله پژوهشی

تاریخ دریافت: ۱۴۰۵/۰۶/۱۶

تاریخ بازنگری: ۱۴۰۵/۰۷/۱۷

تاریخ پذیرش: ۱۴۰۵/۰۸/۲۱

تاریخ انتشار: ۱۴۰۵/۰۸/۰۱

کلیدواژه ها:

سیستم اطلاعات جغرافیایی

نقاط پرخطر شهری تحلیل مکانی هوش مصنوعی مدیریت شهری

چکیده

افزایش جمعیت شهری، توسعه نامتوازن کالبدی و رشد حجم ترددهای درون‌شهری، ضرورت شناسایی و پیش‌بینی نقاط پرخطر شهری را به‌عنوان یکی از مهم‌ترین الزامات مدیریت ایمنی و کاهش آسیب‌پذیری شهری مطرح کرده است. هدف پژوهش حاضر، ارائه یک مدل هوشمند مبتنی بر سیستم اطلاعات جغرافیایی برای شناسایی و پیش‌بینی نقاط پرخطر شهری با تلفیق داده‌های مکانی، محیطی و اجتماعی با الگوریتم‌های یادگیری ماشین است. در این پژوهش، مجموعه‌ای از متغیرهای مؤثر بر شکل‌گیری نقاط پرخطر، شامل تراکم جمعیت، شبکه معابر، شدت و نوع کاربری اراضی، فاصله از مراکز پرتردد، وضعیت دسترسی به خدمات شهری، ویژگی‌های هندسی معابر، سوابق رخدادهای خطرناک و سایر شاخص‌های کالبدی و محیطی، در محیط GIS گردآوری، استانداردسازی و تحلیل می‌شوند. سپس با استفاده از الگوریتم‌های یادگیری ماشین، از جمله Random Forest، Support Vector Machine و XGBoost، الگوی فضایی نقاط پرخطر استخراج و احتمال وقوع خطر در محدوده‌های مختلف شهری پیش‌بینی می‌شود. به‌منظور افزایش دقت مدل، اهمیت نسبی متغیرها با استفاده از روش‌های تفسیرپذیری مدل ارزیابی شده و عملکرد الگوریتم‌ها بر اساس شاخص‌هایی نظیر Accuracy، Precision، Recall، F1-Score و AUC-ROC مقایسه خواهد شد. در نهایت، خروجی مدل در قالب نقشه پهنه‌بندی خطر در محیط GIS ارائه می‌شود تا امکان شناسایی نواحی دارای ریسک بالا و اولویت‌بندی مداخلات مدیریتی و زیرساختی فراهم شود.

استناد: رحیمیان، علیرضا. (۱۴۰۵). ارائه مدل هوشمند GIS برای شناسایی و پیش‌بینی نقاط پرخطر شهری با تلفیق داده‌های مکانی و یادگیری ماشین. جغرافیا و برنامه ریزی منطقه‌ای، ۱۶(۴)، ۲۸-۴۴. DOI: 10.22034/jgeoq.2026.601292.4542

۱. مقدمه

رشد شتابان شهرنشینی، افزایش تراکم جمعیت و توسعه پیچیده شبکه‌های حمل‌ونقل، کاربری‌های شهری و زیرساخت‌های حیاتی، شهرها را با مجموعه‌ای از مخاطرات و تهدیدهای فضایی و محیطی مواجه کرده است. در چنین شرایطی، شناسایی محدوده‌های پرخطر و پیش‌بینی احتمال وقوع رخدادهای نامطلوب، یکی از الزامات اساسی در برنامه‌ریزی شهری، مدیریت بحران و افزایش تاب‌آوری شهرها محسوب می‌شود. نقاط پرخطر شهری می‌توانند تحت تأثیر عوامل متعددی از جمله تراکم جمعیت، الگوی کاربری زمین، ویژگی‌های شبکه معابر، شدت تردد، وضعیت زیرساخت‌ها، شرایط محیطی، دسترسی به خدمات اضطراری و سابقه رخدادهای خطرناک شکل بگیرند. از این رو، تحلیل این نقاط صرفاً بر مبنای یک عامل منفرد، قادر به نمایش کامل ماهیت پیچیده و چندبعدی خطر در محیط شهری نیست (۱، ۲).

در دهه‌های اخیر، سیستم اطلاعات جغرافیایی به یکی از مهم‌ترین ابزارهای تحلیل و مدیریت اطلاعات مکانی در حوزه برنامه‌ریزی شهری و مدیریت خطر تبدیل شده است. قابلیت GIS در ذخیره‌سازی، پردازش، ترکیب و نمایش داده‌های مکانی، امکان بررسی هم‌زمان عوامل مختلف مؤثر بر خطر و تولید نقشه‌های پهنه‌بندی را فراهم می‌کند. مطالعات مختلف نشان داده‌اند که GIS می‌تواند به‌عنوان یک ابزار تصمیم‌یار مؤثر برای ارزیابی خطر، مدیریت دارایی‌های شهری، تحلیل آسیب‌پذیری و افزایش تاب‌آوری مورد استفاده قرار گیرد (۲، ۳). با وجود این، روش‌های سنتی GIS که عمدتاً بر تحلیل‌های توصیفی، شاخص‌های مکانی یا روش‌های وزن‌دهی مبتنی هستند، در مواجهه با روابط غیرخطی، پیچیده و چندعاملی میان متغیرهای شهری با محدودیت مواجه می‌شوند.

یکی از تحولات مهم در این زمینه، ظهور هوش مصنوعی و به‌ویژه الگوریتم‌های یادگیری ماشین در تحلیل داده‌های مکانی است. یادگیری ماشین این امکان را فراهم می‌کند که روابط پنهان و الگوهای پیچیده میان متغیرهای مستقل و وقوع رخدادهای خطرناک بدون تعریف صریح تمام روابط توسط پژوهشگر استخراج شود. این ویژگی در محیط‌های شهری که رفتار سیستم‌ها تحت تأثیر مجموعه‌ای از عوامل کالبدی، اجتماعی، اقتصادی، محیطی و زمانی قرار دارد، اهمیت ویژه‌ای پیدا می‌کند. در سال‌های اخیر، ترکیب GIS و یادگیری ماشین به‌عنوان یکی از رویکردهای مهم در توسعه مدل‌های پیش‌بینی مکانی و ارزیابی خطر مطرح شده است (۴، ۵).

مرور پژوهش‌های جدید نشان می‌دهد که مدل‌های یادگیری ماشین و یادگیری عمیق به‌طور فزاینده‌ای برای تهیه نقشه‌های حساسیت و خطر مورد استفاده قرار گرفته‌اند. این مدل‌ها به‌ویژه در شرایطی که تعداد متغیرهای ورودی زیاد و روابط میان آن‌ها غیرخطی باشد، می‌توانند عملکرد مناسبی ارائه دهند. یک مرور ادبیات در سال ۲۰۲۴ نشان داد که استفاده از الگوریتم‌های یادگیری ماشین و یادگیری عمیق برای مدل‌سازی حساسیت مخاطرات به‌طور محسوسی گسترش یافته و ترکیب داده‌های مکانی و مدل‌های هوشمند می‌تواند ظرفیت پیش‌بینی و آمادگی در برابر مخاطرات را افزایش دهد (۵). همچنین، مطالعات جدید بر اهمیت استفاده از داده‌های چندمنبعی و تلفیق اطلاعات مکانی، محیطی و کاربری زمین برای افزایش دقت مدل‌های خطر تأکید کرده‌اند (۶).

در میان الگوریتم‌های یادگیری ماشین، روش‌هایی نظیر جنگل تصادفی، ماشین بردار پشتیبان، درخت تصمیم و الگوریتم‌های تقویتی نظیر XGBoost به دلیل توانایی در مدل‌سازی روابط غیرخطی و پردازش تعداد زیادی از متغیرهای توضیحی، کاربرد گسترده‌ای در مسائل مکانی پیدا کرده‌اند. به‌عنوان نمونه، پژوهشی درباره شناسایی و پیش‌بینی نقاط پرخطر تصادفات در شبکه معابر شهری، با تلفیق داده‌های مکانی–زمانی و یادگیری ماشین از XGBoost برای طبقه‌بندی نقاط به نواحی پرخطر، عادی و کم‌خطر استفاده کرده است. نتایج این پژوهش نشان داد که در نظر گرفتن ویژگی‌های نقاط مجاور و روابط مکانی–زمانی می‌تواند دقت پیش‌بینی نقاط پرخطر را افزایش داده و نرخ شناسایی اشتباه را کاهش دهد (۷). این یافته اهمیت در نظر گرفتن «وابستگی مکانی» را در مدل‌های پیش‌بینی خطر شهری نشان می‌دهد.

از سوی دیگر، استفاده از روش‌های آماری فضایی در کنار الگوریتم‌های یادگیری ماشین می‌تواند امکان شناسایی دقیق‌تر خوشه‌های خطر را فراهم کند. شاخص‌هایی مانند Moran's I و Getis-Ord Gi* قادرند الگوهای خوشه‌ای و تمرکز فضایی رخدادها را شناسایی کنند، در حالی که الگوریتم‌های یادگیری ماشین می‌توانند احتمال وقوع خطر را بر اساس مجموعه‌ای از متغیرهای پیش‌بین برآورد کنند. مطالعات انجام‌شده در زمینه نقاط پرخطر تصادفات نشان داده‌اند که ترکیب روش‌های آماری فضایی و

Random Forest می‌تواند برای شناسایی و پیش‌بینی نقاط پرخطر مورد استفاده قرار گیرد (۸). بنابراین، تلفیق دو رویکرد «تحلیل فضایی» و «یادگیری پیش‌بینانه» می‌تواند نسبت به استفاده منفرد از هر یک، چارچوب جامع‌تری برای تحلیل خطر شهری ایجاد کند.

اهمیت این موضوع زمانی بیشتر می‌شود که خطر شهری به‌عنوان پدیده‌ای ایستا در نظر گرفته نشود. بسیاری از مخاطرات شهری دارای ماهیت زمانی و مکانی هم‌زمان هستند؛ به این معنا که احتمال وقوع آن‌ها در مکان‌های مختلف و در ساعات، روزها یا دوره‌های زمانی متفاوت تغییر می‌کند. برای مثال، الگوی خطر در شبکه معابر می‌تواند تحت تأثیر حجم ترافیک، فعالیت‌های تجاری، دسترسی به مراکز حمل‌ونقل عمومی و ویژگی‌های کاربری زمین در طول زمان تغییر کند. بنابراین، مدل‌های پیش‌بینی نوین باید علاوه بر ویژگی‌های مکانی، قابلیت استفاده از اطلاعات زمانی و ویژگی‌های محیط پیرامون هر نقطه را نیز داشته باشند (۷).

یکی دیگر از چالش‌های مهم مدل‌های هوشمند مکانی، مسئله تفسیرپذیری است. اگرچه مدل‌های پیچیده یادگیری ماشین می‌توانند دقت پیش‌بینی بالایی داشته باشند، اما در بسیاری از موارد مشخص نیست که کدام متغیرها بیشترین نقش را در پیش‌بینی خطر داشته‌اند. این مسئله در مدیریت شهری اهمیت ویژه‌ای دارد؛ زیرا تصمیم‌گیران علاوه بر دریافت نقشه خطر، نیازمند شناخت عوامل مؤثر بر ایجاد آن هستند. در سال‌های اخیر، استفاده از روش‌هایی مانند SHAP برای تفسیر خروجی مدل‌های یادگیری ماشین و تعیین اهمیت متغیرها در مطالعات مکانی مورد توجه قرار گرفته است (۷). همچنین، پژوهش‌های جدید در حوزه GeoAI بر ضرورت توسعه مدل‌های مسئولانه، شفاف و قابل تفسیر برای کاربردهای شهری تأکید دارند (۹).

با وجود پیشرفت‌های قابل توجه، همچنان شکاف پژوهشی مهمی در زمینه ارائه مدل یکپارچه برای شناسایی و پیش‌بینی نقاط پرخطر شهری وجود دارد. بخش قابل توجهی از مطالعات بر یک نوع مخاطره خاص، مانند سیلاب، زمین‌لغزش یا تصادفات تمرکز کرده‌اند و کمتر به توسعه یک چارچوب عمومی و قابل انطباق برای ترکیب انواع داده‌های مکانی، محیطی، کالبدی و اجتماعی پرداخته‌اند. همچنین، کیفیت داده، ناهمگونی مکانی، وابستگی فضایی، عدم توازن داده‌های رخداد و غیررخداد، انتخاب ویژگی‌های مناسب و تفسیرپذیری مدل از جمله چالش‌هایی هستند که می‌توانند بر عملکرد مدل‌های هوشمند مکانی اثرگذار باشند (۴، ۵، ۹).

در این راستا، پژوهش حاضر با هدف ارائه یک مدل هوشمند GIS برای شناسایی و پیش‌بینی نقاط پرخطر شهری، بر تلفیق قابلیت‌های تحلیل مکانی GIS با الگوریتم‌های یادگیری ماشین تمرکز دارد. در چارچوب پیشنهادی، ابتدا داده‌های مکانی مرتبط با ویژگی‌های کالبدی، جمعیتی، کاربری زمین، شبکه معابر، دسترسی به خدمات و سوابق رخدادهای خطرناک گردآوری و در محیط GIS یکپارچه می‌شوند. سپس با استفاده از روش‌های تحلیل فضایی، الگوهای مکانی و خوشه‌های خطر شناسایی شده و متغیرهای مؤثر به‌عنوان ورودی مدل‌های یادگیری ماشین مورد استفاده قرار می‌گیرند. در مرحله بعد، الگوریتم‌های منتخب برای پیش‌بینی احتمال قرارگیری هر محدوده در طبقات مختلف خطر آموزش داده شده و عملکرد آن‌ها با معیارهای استاندارد ارزیابی خواهد شد. نوآوری اصلی پژوهش در ایجاد پیوند میان تحلیل فضایی، داده‌های چندمنبعی و یادگیری ماشین با هدف عبور از شناسایی صرفِ نقاط پرخطر و حرکت به سمت پیش‌بینی مکانی خطر است. چنین رویکردی می‌تواند از یک سو دقت تحلیل‌های مکانی را افزایش دهد و از سوی دیگر، اطلاعات قابل استفاده‌تری برای تصمیم‌گیری در اختیار مدیران شهری قرار دهد. خروجی نهایی مدل نیز می‌تواند در قالب نقشه هوشمند پهنه‌بندی خطر ارائه شود تا نواحی دارای ریسک بالاتر، عوامل مؤثر بر آن‌ها و اولویت‌های مداخله مشخص شوند. در نهایت، انتظار می‌رود چارچوب پیشنهادی بتواند به‌عنوان ابزاری برای مدیریت پیشگیرانه، تخصیص هدفمند منابع، بهبود برنامه‌ریزی زیرساختی و ارتقای تاب‌آوری شهری مورد استفاده قرار گیرد (۲، ۳، ۵).

۲. مبانی نظری و پیشینه پژوهش

۲-۱. مبانی نظری پژوهش

۲-۱-۱. نقاط پرخطر شهری

نقطه پرخطر شهری به محدوده‌ای اطلاق می‌شود که احتمال وقوع یک یا چند نوع رخداد نامطلوب در آن، در مقایسه با سایر نواحی شهری بیشتر است. مفهوم نقاط پرخطر در مطالعات شهری معمولاً با تمرکز فضایی رخدادهایی مانند تصادفات، جرائم، حوادث، آتش‌سوزی، سیلاب و سایر مخاطرات ارتباط دارد. مهم‌ترین ویژگی این نقاط، وجود نوعی «خوشه‌بندی فضایی» است؛ به

این معنا که رخدادهای خطرناک در سطح شهر به‌صورت تصادفی توزیع نشده و در برخی محدوده‌ها تمرکز بیشتری پیدا می‌کنند. از این منظر، شناسایی نقاط پرخطر نیازمند توجه هم‌زمان به موقعیت مکانی رخدادها، ویژگی‌های محیط پیرامون و روابط میان واحدهای فضایی است.

در مطالعات سنتی، شناسایی نقاط پرخطر عمدتاً با استفاده از روش‌های آماری فضایی نظیر شاخص Moran's I، آماره Getis-Ord Gi* و Kernel Density Estimation انجام شده است. این روش‌ها برای تشخیص خوشه‌های فضایی و تعیین نواحی دارای تمرکز بالای رخدادها بسیار مفید هستند، اما الزاماً قادر به پیش‌بینی رفتار آینده نقاط خطر یا مدل‌سازی روابط پیچیده میان تعداد زیادی از عوامل نیستند. برای نمونه، آتوموآ و همکاران با استفاده از آمار فضایی و Random Forest نشان دادند که ترکیب روش‌های شناسایی خوشه‌های فضایی با یادگیری ماشین می‌تواند برای شناسایی و پیش‌بینی نقاط پرخطر تصادفات مورد استفاده قرار گیرد؛ مدل Random Forest در این مطالعه به دقت ۷۶/۷ درصد در اعتبارسنجی و ۷۴ درصد در پیش‌بینی دست یافت (۱).

۲-۱-۲. سیستم اطلاعات جغرافیایی و تحلیل مکانی

سیستم اطلاعات جغرافیایی، چارچوبی برای گردآوری، ذخیره‌سازی، مدیریت، تحلیل و نمایش داده‌هایی است که دارای مؤلفه مکانی هستند. ویژگی اصلی GIS در مطالعات خطر شهری، قابلیت تلفیق داده‌های ناهمگون در یک محیط تحلیلی واحد است. برای مثال، اطلاعات مربوط به شبکه معابر، کاربری زمین، تراکم جمعیت، مراکز خدماتی، زیرساخت‌ها، وضعیت توپوگرافی و سوابق رخدادهای خطرناک را می‌توان در GIS بر روی یکدیگر قرار داد و روابط فضایی آن‌ها را تحلیل کرد.

از دیدگاه نظری، تحلیل خطر شهری در محیط GIS مبتنی بر این فرض است که وقوع یک رخداد خطرناک تابعی از ویژگی‌های مکانی و محیط پیرامون آن است. بنابراین می‌توان خطر را به‌صورت تابعی از مجموعه متغیرهای فضایی تعریف کرد:

$$Risk = f(X1, X2, \dots, Xn)$$

که در آن Xi متغیرهای مؤثر بر خطر، مانند تراکم جمعیت، کاربری زمین، فاصله از معابر اصلی، تراکم تقاطع‌ها، دسترسی به خدمات و سوابق رخدادها هستند. GIS امکان تولید این متغیرها، استانداردسازی آن‌ها و تبدیل اطلاعات خام مکانی به متغیرهای قابل استفاده در مدل‌های پیش‌بینی را فراهم می‌کند.

۳-۱-۲. یادگیری ماشین در تحلیل خطر شهری

یادگیری ماشین مجموعه‌ای از روش‌های محاسباتی است که به سیستم اجازه می‌دهد الگوهای موجود در داده‌ها را شناسایی کرده و بر اساس آن‌ها برای داده‌های جدید پیش‌بینی انجام دهد. در مدل‌سازی خطر شهری، یادگیری ماشین می‌تواند روابط غیرخطی میان متغیرهای محیطی، کالبدی، اجتماعی و مکانی را شناسایی کند؛ روابطی که شناسایی آن‌ها با مدل‌های خطی یا روش‌های وزن‌دهی سنتی دشوار است.

الگوریتم‌هایی مانند Random Forest، Support Vector Machine، XGBoost و شبکه‌های عصبی از جمله روش‌هایی هستند که در مطالعات مکانی و شهری کاربرد پیدا کرده‌اند. مزیت مهم این الگوریتم‌ها، توانایی پردازش تعداد زیادی متغیر و شناسایی الگوهای پیچیده است. با این حال، عملکرد مدل به کیفیت داده، انتخاب متغیرهای مناسب، نحوه نمونه‌برداری فضایی و چگونگی تفکیک داده‌های آموزش و آزمون وابسته است.

٤-۱-۲.GeoAI و تلفیق GIS با هوش مصنوعی

ترکیب هوش مصنوعی و فناوری‌های مکانی در سال‌های اخیر با مفهوم هوش مصنوعی مکانی یا GeoAI شناخته می‌شود. GeoAI را می‌توان رویکردی دانست که قابلیت‌های تحلیل فضایی، داده‌های جغرافیایی و الگوریتم‌های هوش مصنوعی را در یک چارچوب واحد ترکیب می‌کند. این رویکرد امکان عبور از تحلیل‌های صرفاً توصیفی و حرکت به سمت تحلیل‌های پیش‌بینانه و تصمیم‌یار را فراهم می‌سازد.

ماراسینگه و همکاران در یک مطالعه مروری در سال ۲۰۲۴ نشان دادند که کاربرد GeoAI در برنامه‌ریزی و تصمیم‌گیری شهری به سرعت در حال گسترش است و تلفیق هوش مصنوعی با تحلیل‌های مکانی می‌تواند ظرفیت تحلیل شهری را افزایش دهد. این

مطالعه همچنین بر اهمیت تفسیرپذیری، شفافیت، حاکمیت داده، عدالت، حریم خصوصی و مسئولیت‌پذیری در کاربردهای شهری GeoAI تأکید کرده است (۲).

۵-۱-۲. نقش روابط مکانی و زمانی

یکی از اصول اساسی در تحلیل نقاط پرخطر این است که رخدادهای شهری از محیط پیرامون خود مستقل نیستند. ویژگی‌های یک نقطه می‌تواند تحت تأثیر ویژگی‌های نقاط مجاور قرار گیرد. بنابراین، استفاده از متغیرهای مربوط به همسایگی فضایی می‌تواند توان پیش‌بینی مدل را افزایش دهد.

وو و همکاران در سال ۲۰۲۳ با تلفیق Space-Time Cube و یادگیری ماشین، نقاط پرخطر تصادفات را در شبکه معابر شهری منهتن بررسی کردند. آنان از XGBoost برای طبقه‌بندی نقاط به سه گروه پرخطر، عادی و کم‌خطر استفاده کردند و نشان دادند که ویژگی‌های نقاط همسایه و روابط مکانی – زمانی نقش مهمی در پیش‌بینی خطر دارند. همچنین استفاده از SHAP امکان تفسیر اثر متغیرهای مرتبط با کاربری زمین، حمل‌ونقل عمومی و امکانات شهری را فراهم کرد (۳).

۲-۲. پیشینه پژوهش

مطالعات انجام‌شده نشان می‌دهد روند پژوهش در زمینه شناسایی نقاط پرخطر از روش‌های آماری فضایی به سمت مدل‌های هوشمند و ترکیبی حرکت کرده است. در مطالعات اولیه، تمرکز عمدتاً بر شناسایی خوشه‌های فضایی بود، در حالی که پژوهش‌های جدیدتر تلاش می‌کنند علاوه بر شناسایی وضعیت موجود، احتمال وقوع خطر در آینده را نیز پیش‌بینی کنند.

ردیف	پژوهشگر / سال	روش و داده	مهم‌ترین یافته	ارتباط با پژوهش حاضر
۱	آتوموآ و همکاران			
(۲۰۲۲)	Getis-Ord + Random Forest	شناسایی و پیش‌بینی نقاط پرخطر تصادفات؛ دقت ۷۶/۷٪ در اعتبارسنجی و ۷۴٪ در پیش‌بینی	تأیید قابلیت تلفیق آمار فضایی و ML (1)	
۲	وو و همکاران			
(۲۰۲۳)	Space-Time Cube + XGBoost + SHAP	افزایش دقت شناسایی نقاط پرخطر با توجه به ویژگی نقاط همسایه و روابط مکانی–زمانی	مبنای استفاده از ویژگی‌های مکانی و زمانی (۳)	
۳	ماراسینگه و همکاران			
(۲۰۲۴)	مرور نظام‌مند GeoAI	شناسایی ظرفیت‌ها و چالش‌های استفاده از GeoAI در برنامه‌ریزی شهری	مبنای نظری تلفیق GIS و AI (2)	
۴	اکیانیو و همکاران			
(۲۰۲۲)	GIS + Machine Learning	بررسی قابلیت تلفیق GIS و یادگیری ماشین برای تحلیل مسائل مکانی	تأیید چارچوب ترکیبی GIS-ML	

۵	مطالعات مختلف حوزه تحلیل خطر شهری	GIS + روش‌های آماری فضایی	شناسایی خوشه‌ها و الگوهای تمرکز خطر	مبنای مرحله تحلیل فضایی
۶	مطالعات یادگیری ماشین در ایمنی شهری	Random Forest / XGBoost	استخراج روابط غیرخطی بین ویژگی‌های محیطی و خطر	مبنای مرحله پیش‌بینی مدل

نتایج این مطالعات نشان می‌دهد که استفاده منفرد از GIS اگرچه در شناسایی الگوهای فضایی بسیار مؤثر است، اما برای پیش‌بینی دقیق رفتار آینده خطر محدودیت دارد. از طرف دیگر، استفاده منفرد از یادگیری ماشین ممکن است ساختار فضایی داده‌ها را نادیده بگیرد. بنابراین، ترکیب این دو رویکرد می‌تواند مزیت‌های هر دو روش را هم‌زمان مورد استفاده قرار دهد.

۳-۲. جمع‌بندی مبانی نظری و شکاف پژوهش

بررسی مبانی نظری و مطالعات پیشین نشان می‌دهد سه جریان اصلی در پژوهش‌های مرتبط وجود دارد: نخست، روش‌های تحلیل فضایی برای شناسایی خوشه‌های خطر؛ دوم، روش‌های یادگیری ماشین برای پیش‌بینی رخداد؛ و سوم، رویکردهای جدید GeoAI برای ترکیب داده‌های مکانی و هوش مصنوعی. با وجود پیشرفت‌های قابل توجه، همچنان نیاز به مدل‌هایی وجود دارد که بتوانند داده‌های چندمنبعی مکانی را در GIS یکپارچه کرده، روابط فضایی و در صورت امکان زمانی را لحاظ کنند، با الگوریتم‌های یادگیری ماشین خطر را پیش‌بینی کنند و در نهایت عوامل مؤثر بر پیش‌بینی را تفسیر نمایند.

بر همین اساس، پژوهش حاضر درصدد ارائه مدلی یکپارچه است که GIS را به‌عنوان بستر مدیریت و تحلیل داده‌های مکانی و یادگیری ماشین را به‌عنوان موتور پیش‌بینی به کار گیرد. تفاوت اصلی این رویکرد با مطالعات سنتی، عبور از صرفاً «نقشه‌برداری نقاط پرخطر» و حرکت به سمت «پیش‌بینی هوشمند نقاط مستعد خطر» است. همچنین استفاده از روش‌های تفسیرپذیری می‌تواند مشخص کند کدام عوامل بیشترین نقش را در افزایش احتمال قرارگیری یک محدوده در طبقه خطر بالا دارند.

۳. روش پژوهش

پژوهش حاضر از نظر هدف، کاربردی و از نظر روش، توصیفی – تحلیلی و مدل‌سازی است و با هدف ارائه مدل هوشمند GIS برای شناسایی و پیش‌بینی نقاط پرخطر شهری انجام می‌شود. جامعه آماری پژوهش شامل کلیه رخدادهای ثبت‌شده مرتبط با نقاط پرخطر در محدوده مورد مطالعه طی یک دوره زمانی مشخص است. داده‌های مورد نیاز شامل موقعیت مکانی رخدادها، شبکه معابر، تراکم جمعیت، کاربری اراضی، تراکم تقاطع‌ها، فاصله از مراکز خدماتی و درمانی، ایستگاه‌های حمل‌ونقل عمومی، ویژگی‌های کالبدی و سایر متغیرهای مؤثر بر خطر است که از منابع رسمی و پایگاه‌های مکانی گردآوری می‌شوند.

در مرحله نخست، داده‌ها در محیط GIS پیش‌پردازش، هم‌مرجع و یکپارچه شده و شاخص‌های مکانی استخراج می‌شوند. سپس با استفاده از روش‌های تحلیل فضایی نظیر Kernel Density و Getis-Ord Gi، الگوهای تمرکز و خوشه‌های پرخطر شناسایی می‌شوند. در مرحله بعد، متغیرهای منتخب به‌عنوان ورودی الگوریتم‌های Random Forest، Support Vector Machine و XGBoost به کار گرفته می‌شوند. داده‌ها به مجموعه‌های آموزش و آزمون تقسیم شده و عملکرد مدل‌ها با معیارهای Accuracy، Precision، Recall، F1-Score و AUC ارزیابی می‌شود. در نهایت، بهترین مدل انتخاب و خروجی آن در محیط GIS به‌صورت نقشه پیش‌بینی و پهنه‌بندی نقاط پرخطر ارائه خواهد شد. همچنین با استفاده از روش SHAP، اهمیت متغیرهای مؤثر در پیش‌بینی خطر تفسیر می‌شود.

٤. نتایج و تحلیل

٤-۱. نتایج حاصل از آماده‌سازی و تحلیل داده‌های مکانی

در این پژوهش، پس از گردآوری داده‌های مکانی مرتبط با رخدادهای پرخطر شهری، تمامی لایه‌ها در محیط سیستم اطلاعات جغرافیایی یکپارچه و برای ورود به مدل‌های یادگیری ماشین آماده شدند. مهم‌ترین لایه‌های مورد استفاده شامل شبکه معابر، تراکم جمعیت، کاربری اراضی، تراکم تقاطع‌ها، فاصله از مراکز خدماتی، فاصله از مراکز پرتردد، ایستگاه‌های حمل‌ونقل عمومی، ویژگی‌های هندسی معابر و سابقه رخدادهای خطرناک بود.

به‌منظور حذف اثر مقیاس متفاوت متغیرها، متغیرهای کمی استانداردسازی شدند. همچنین داده‌های دارای مختصات ناقص یا اطلاعات غیرقابل استفاده حذف و داده‌های مکانی بر اساس یک سیستم مختصات یکسان به محیط GIS منتقل شدند. سپس محدوده مورد مطالعه به واحدهای مکانی منظم تقسیم شد تا برای هر واحد، مجموعه‌ای از ویژگی‌های مکانی و سابقه رخدادها محاسبه شود.

نتایج اولیه تحلیل فضایی نشان داد که رخدادهای خطرناک از الگوی کاملاً تصادفی پیروی نمی‌کنند و در برخی محدوده‌های شهری تمرکز بیشتری دارند. این نتیجه از نظر نظری با مطالعات پیشین همخوان است؛ برای مثال، آتوموآ و همکاران نشان دادند که رخدادهای تصادفات دارای وابستگی فضایی بوده و استفاده از آماره‌های فضایی می‌تواند برای شناسایی خوشه‌های پرخطر مؤثر باشد.

۴-۲. شناسایی خوشه‌های پرخطر با تحلیل فضایی

در مرحله بعد، به‌منظور شناسایی مناطق دارای تمرکز بالای رخدادها، از روش‌های Kernel Density و Getis-Ord Gi استفاده شد. خروجی تحلیل Kernel Density نشان داد که شدت رخدادها در سطح شهر یکنواخت نیست و چند محدوده دارای تراکم بالاتری از رخدادها هستند.

در تحلیل Getis-Ord Gi، نواحی شهری به سه گروه کلی شامل نقاط پرخطر، نقاط عادی و نقاط کم‌خطر تقسیم شدند. نقاط دارای مقادیر مثبت و معنادار آماره Gi* به‌عنوان خوشه‌های پرخطر و نقاط دارای مقادیر منفی و معنادار به‌عنوان خوشه‌های کم‌خطر شناسایی شدند.

در سناریوی نمونه پژوهش، حدود ۱۴ درصد از واحدهای مکانی در طبقه پرخطر، ۶۱ درصد در طبقه خطر متوسط و ۲۵ درصد در طبقه کم‌خطر قرار گرفتند. این یافته نشان می‌دهد که بخش نسبتاً محدودی از سطح شهری می‌تواند سهم قابل توجهی از رخدادهای خطرناک را به خود اختصاص دهد.

جدول ۱. توزیع نمونه واحدهای مکانی بر اساس سطح خطر

تفسیر	درصد واحدهای مکانی	طبقه خطر
احتمال پایین وقوع رخداد	۲۵٪	کم‌خطر
نیازمند پایش و مدیریت مستمر	۶۱٪	متوسط
اولویت نخست مداخلات شهری	۱۴٪	پرخطر
–	۱۰۰٪	مجموع

الگوی فوق نشان می‌دهد که مدیریت خطر شهری نباید منابع را به‌صورت یکنواخت در کل محدوده شهر توزیع کند؛ بلکه تمرکز منابع بر خوشه‌های پرخطر می‌تواند کارایی بیشتری داشته باشد.

۴-۳. تحلیل اهمیت متغیرهای مکانی

پس از تولید لایه‌های مکانی، اهمیت نسبی متغیرها در مدل یادگیری ماشین بررسی شد. نتایج نمونه نشان داد که تراکم رخدادهای قبلی، تراکم تقاطع‌ها، شدت تردد، تراکم کاربری‌های تجاری و فاصله از مراکز حمل‌ونقل عمومی از مهم‌ترین عوامل مؤثر بر افزایش احتمال قرارگیری یک محدوده در طبقه پرخطر بودند.

در مقابل، برخی متغیرها مانند فاصله از مراکز خدماتی یا برخی کاربری‌های کم‌تراکم، اهمیت کمتری در پیش‌بینی خطر داشتند. این نتیجه نشان می‌دهد که همه متغیرهای مکانی وزن یکسانی در مدل ندارند و الگوریتم یادگیری ماشین می‌تواند اهمیت نسبی آن‌ها را از طریق الگوهای موجود در داده‌ها استخراج کند.

جدول ۲. اهمیت نسبی نمونه متغیرهای ورودی مدل

رتبه	متغیر	اهمیت نسبی نمونه
۱	تراکم رخدادهای قبلی	۱۸٪/۶
۲	تراکم تقاطع‌ها	۱۸٪/۴
۳	شدت تردد	۱۳٪/۷
۴	تراکم کاربری تجاری	۱۱٪/۹
۵	فاصله از حمل‌ونقل عمومی	۱۰٪/۴
۶	تراکم جمعیت	۹٪/۸
۷	ویژگی شبکه معابر	۸٪/۱
۸	فاصله از مراکز خدماتی	۶٪/۷
۹	سایر متغیرها	۵٪/۹
مجموع	–	۱۰۰٪

بر اساس این نتایج، می‌توان استدلال کرد که ویژگی‌های محیط پیرامون و سابقه مکانی رخدادها نقش مهمی در پیش‌بینی نقاط پرخطر دارند. این موضوع با یافته‌های وو و همکاران نیز سازگار است؛ آنان در مطالعه خود نشان دادند که ویژگی‌های واحدهای مکانی مجاور نقش مهمی در پیش‌بینی خطر داشته و بسیاری از متغیرهای کلیدی از ویژگی‌های نقاط همسایه استخراج شده‌اند.

٤-٤. مقایسه عملکرد الگوریتم‌های یادگیری ماشین

در مرحله مدل‌سازی، سه الگوریتم Random Forest، Support Vector Machine و XGBoost برای طبقه‌بندی واحدهای مکانی بر اساس سطح خطر مورد استفاده قرار گرفتند. داده‌ها به دو بخش آموزش و آزمون تقسیم شدند. سپس عملکرد هر الگوریتم با استفاده از Accuracy، Precision، Recall، F1-Score و AUC-ROC ارزیابی شد.

در نتایج نمونه، XGBoost بهترین عملکرد کلی را نشان داد و پس از آن Random Forest و Support Vector Machine قرار گرفتند.

جدول ۳: مقایسه عملکرد نمونه الگوریتم‌های یادگیری ماشین

جدول ۳. مقایسه عملکرد الگوریتم‌های یادگیری ماشین

مدل	Accuracy	Precision	Recall	F1-Score	AUC
Random Forest	۰.۸۶	۰.۸۴	۰.۸۲	۰.۸۳	۰.۹۰
SVM	۰.۸۱	۰.۷۹	۰.۷۷	۰.۷۸	۰.۸۶
XGBoost	۰.۸۹	۰.۸۸	۰.۸۶	۰.۸۷	۰.۹۳

بر اساس جدول، مدل XGBoost با Accuracy برابر با ۰/۸۹، Precision برابر با ۰/۸۸، Recall برابر با ۰/۸۶ و F1-Score برابر با ۰/۸۷ بهترین عملکرد را در میان مدل‌های بررسی‌شده نشان می‌دهد. همچنین مقدار AUC برابر با ۰/۹۳ بیانگر قدرت تفکیک مناسب مدل بین واحدهای پرخطر و غیرپرخطر است.

برتری XGBoost را می‌توان به توانایی آن در مدل‌سازی روابط غیرخطی و تعامل میان متغیرها نسبت داد. این موضوع با پژوهش وو و همکاران مطابقت دارد که از XGBoost برای طبقه‌بندی الگوهای خطر در واحدهای مکانی و زمانی استفاده کردند و عملکرد مناسب این الگوریتم را در مقایسه با روش‌های متعارف گزارش کردند.

٤-۵. تحلیل ماتریس خطا

برای ارزیابی دقیق‌تر مدل منتخب، ماتریس خطا مورد بررسی قرار گرفت. در یک مدل طبقه‌بندی سه‌گانه، خطاهای اصلی شامل اشتباه گرفتن نقاط پرخطر با نقاط متوسط یا کم‌خطر و بالعکس است.

در کاربردهای مدیریت شهری، خطای نوع دوم یعنی شناسایی نکردن یک نقطه واقعاً پرخطر اهمیت ویژه‌ای دارد. زیرا ممکن است باعث شود مدیران شهری از یک منطقه دارای ریسک بالا غافل شوند. بنابراین، صرفاً افزایش Accuracy کافی نیست و باید Recall مربوط به طبقه پرخطر نیز مورد توجه قرار گیرد.

در سناریوی نمونه، Recall طبقه پرخطر برابر با ۰/۸۶ به دست آمد. این مقدار نشان می‌دهد که مدل توانسته است بخش عمده واحدهای واقعاً پرخطر را شناسایی کند. در عین حال، وجود تعدادی خطای طبقه‌بندی نشان می‌دهد که مدل هنوز نیازمند بهبود داده‌ها و تنظیم پارامترهای الگوریتم است.

٦-٤. تحلیل فضایی خروجی مدل

پس از انتخاب XGBoost به‌عنوان مدل برتر، احتمال خطر برای تمامی واحدهای مکانی محاسبه شد و خروجی مدل به محیط GIS منتقل گردید. بر اساس این خروجی، مناطق شهری در پنج طبقه بسیار کم‌خطر، کم‌خطر، متوسط، پرخطر و بسیار پرخطر طبقه‌بندی شدند.

جدول ٤. نمونه طبقه‌بندی نهایی پهنه‌های خطر

طبقه	دامنه احتمال خطر	ویژگی
بسیار کم‌خطر	۰ تا ۰/۲۰	ریسک بسیار پایین
کم‌خطر	۰/۲۰ تا ۰/۴۰	ریسک پایین
متوسط	۰/۴۰ تا ۰/۶۰	نیازمند پایش
پرخطر	۰/۶۰ تا ۰/۸۰	نیازمند مداخله
بسیار پرخطر	۰/۸۰ تا ۱/۰۰	اولویت فوری

تحلیل فضایی نشان داد که نواحی دارای احتمال خطر بالا عمدتاً در مجاورت محورهای اصلی، تقاطع‌های پرتردد، مناطق دارای تراکم بالای کاربری تجاری و محدوده‌های دارای سابقه بالای رخداد قرار می‌گیرند. این الگو از منظر برنامه‌ریزی شهری منطقی است؛ زیرا افزایش تعاملات انسانی، حجم تردد، تراکم فعالیت‌ها و پیچیدگی شبکه معابر می‌تواند تعداد نقاط تعارض و در نتیجه احتمال رخدادهای خطرناک را افزایش دهد.

٧-٤. نقش همسایگی فضایی در پیش‌بینی خطر

یکی از مهم‌ترین یافته‌های مدل پیشنهادی، اهمیت ویژگی‌های واحدهای مجاور است. تحلیل نمونه نشان داد که یک نقطه شهری را نمی‌توان کاملاً مستقل از محیط اطراف آن ارزیابی کرد. برای مثال، حتی اگر یک قطعه شهری دارای ویژگی‌های نسبتاً مطلوب باشد، قرار گرفتن آن در مجاورت یک محور پرتردد، تقاطع بزرگ یا محدوده دارای تراکم بالای رخداد می‌تواند احتمال خطر آن را افزایش دهد.

این یافته اهمیت استفاده از متغیرهای فضایی مشتق‌شده از همسایگی را نشان می‌دهد. در پژوهش وو و همکاران نیز مشخص شد که بسیاری از ویژگی‌های کلیدی مدل از واحدهای مکانی مجاور ناشی می‌شوند و لحاظ کردن این ویژگی‌ها می‌تواند عملکرد پیش‌بینی را بهبود دهد.

٨-٤. تحلیل تفسیرپذیری مدل با SHAP

برای جلوگیری از تبدیل مدل یادگیری ماشین به یک «جعبه سیاه»، از روش SHAP برای بررسی اثر متغیرها استفاده شد. در این روش، مقدار مثبت SHAP نشان‌دهنده افزایش احتمال قرارگیری یک واحد در طبقه خطر بالاتر و مقدار منفی نشان‌دهنده کاهش این احتمال است.

نتایج نمونه نشان داد که تراکم رخدادهای قبلی و تراکم تقاطع‌ها بیشترین اثر مثبت را بر افزایش احتمال خطر داشته‌اند. شدت تردد و تراکم کاربری تجاری نیز اثر قابل توجهی نشان دادند. در مقابل، افزایش فاصله از مراکز پرتردد و کاهش تراکم رخدادهای تاریخی در برخی شرایط باعث کاهش احتمال خطر شده است.

این قابلیت تفسیرپذیری از نظر مدیریت شهری اهمیت زیادی دارد؛ زیرا مدیران تنها به یک نقشه خطر نیاز ندارند، بلکه باید بدانند چرا یک محدوده در طبقه خطر بالا قرار گرفته است. مطالعه وو و همکاران نیز از SHAP برای تفسیر اثر متغیرهای مربوط به امکانات شهری، حمل‌ونقل عمومی و کاربری زمین استفاده کرده است.

۹-٤. مقایسه با مطالعات پیشین

یافته‌های حاصل از مدل نمونه با نتایج مطالعات پیشین در چند محور اصلی همخوانی دارد. نخست، مطالعه آتوموآ و همکاران نشان داد که ترکیب تحلیل فضایی Getis-Ord با Random Forest می‌تواند برای شناسایی و پیش‌بینی نقاط پرخطر مورد استفاده قرار گیرد. آنان در مطالعه خود Accuracy برابر با ۷۶/۷ درصد برای اعتبارسنجی و ۷۴ درصد برای پیش‌بینی گزارش کردند.

در مقایسه، مدل نمونه حاضر Accuracy برابر ۸۹ درصد را برای XGBoost نشان می‌دهد؛ با این حال، این مقدار صرفاً نمونه و شبیه‌سازی‌شده است و نباید به‌عنوان برتری تجربی قطعی نسبت به مطالعه مذکور تفسیر شود، زیرا داده‌ها، مقیاس مکانی، تعریف خطر و روش اعتبارسنجی دو مطالعه یکسان نیستند.

از سوی دیگر، وو و همکاران با استفاده از Space-Time Cube و XGBoost نشان دادند که وارد کردن اطلاعات مکانی و زمانی نقاط مجاور می‌تواند دقت پیش‌بینی را افزایش داده و نرخ کشف اشتباه را کاهش دهد. این نتیجه اهمیت استفاده از ساختار فضایی و زمانی در مدل پیشنهادی پژوهش حاضر را تأیید می‌کند.

جدول ۵. مقایسه مفهومی پژوهش حاضر با مطالعات منتخب

مطالعه	رویکرد	الگوریتم	ویژگی برجسته
آتوموآ و همکاران (۲۰۲۲)	تحلیل فضایی + ML	Random Forest	شناسایی و پیش‌بینی نقاط پرخطر
وو (۲۰۲۳)	Space-Time Cube + ML	XGBoost	لحاظ کردن همسایگی فضایی–زمانی
پژوهش حاضر	GIS + تحلیل فضایی + ML	RF, SVM, XGBoost	مدل یکپارچه شناسایی و پیش‌بینی
پژوهش حاضر	GIS + SHAP	XGBoost	تفسیر عوامل مؤثر بر خطر

۱۰-٤. تحلیل کاربردی نتایج

از منظر مدیریت شهری، نتایج مدل می‌تواند در سه سطح مورد استفاده قرار گیرد. در سطح نخست، نقشه نقاط پرخطر برای شناسایی محدوده‌های دارای اولویت کاربرد دارد. در سطح دوم، اطلاعات اهمیت متغیرها به مدیران کمک می‌کند تا علت اصلی خطر را شناسایی کنند. برای مثال، اگر یک محدوده عمدتاً به دلیل تراکم تقاطع‌ها در طبقه خطر بالا قرار گرفته باشد، اصلاح هندسی تقاطع، بهبود کنترل ترافیک یا مدیریت جریان عبوری می‌تواند در اولویت قرار گیرد.

در سطح سوم، مدل می‌تواند به سمت پیشگیری حرکت کند. به جای آنکه مدیریت شهری پس از وقوع رخداد اقدام کند، می‌توان از احتمال پیش‌بینی‌شده خطر برای اولویت‌بندی اقدامات قبل از وقوع حادثه استفاده کرد. این تغییر از مدیریت واکنشی به مدیریت پیشگیرانه یکی از مهم‌ترین مزایای استفاده از مدل‌های یادگیری ماشین در محیط GIS است.

۱۱-٤. جمع‌بندی نتایج

به‌طور کلی، نتایج نمونه پژوهش نشان می‌دهد که تلفیق GIS و یادگیری ماشین می‌تواند چارچوب مناسبی برای شناسایی و پیش‌بینی نقاط پرخطر شهری ایجاد کند. تحلیل فضایی امکان شناسایی خوشه‌های خطر و تولید متغیرهای مکانی را فراهم می‌کند، در حالی که الگوریتم‌های یادگیری ماشین قادرند روابط پیچیده و غیرخطی میان این متغیرها را استخراج کنند.

در میان مدل‌های بررسی‌شده، XGBoost در سناریوی نمونه بهترین عملکرد را نشان داد. همچنین تحلیل اهمیت متغیرها نشان داد که سابقه رخداد، تراکم تقاطع‌ها، شدت تردد، کاربری زمین و ویژگی‌های نقاط مجاور می‌توانند نقش مهمی در پیش‌بینی خطر داشته باشند. استفاده از SHAP نیز امکان تفسیر خروجی مدل و تبدیل نتایج محاسباتی به اطلاعات قابل استفاده برای تصمیم‌گیری شهری را فراهم می‌کند.

بنابراین، چارچوب پیشنهادی پژوهش را می‌توان به‌عنوان یک سامانه تصمیم‌یار هوشمند در نظر گرفت که از سه جزء اصلی تشکیل شده است: تحلیل مکانی در GIS، پیش‌بینی با یادگیری ماشین و تفسیرپذیری مدل. خروجی نهایی این سامانه، نقشه‌ای پویا از احتمال خطر است که می‌تواند در برنامه‌ریزی زیرساختی، مدیریت ایمنی، تخصیص منابع و اولویت‌بندی اقدامات پیشگیرانه مورد استفاده قرار گیرد.

با این حال، اعتبار نهایی این نتایج به کیفیت و حجم داده‌های واقعی، تعداد رخدادها، دقت مختصات مکانی، دوره زمانی مطالعه، تعریف دقیق متغیر هدف و نحوه اعتبارسنجی مدل وابسته است. بنابراین در نسخه نهایی مقاله باید اعداد نمونه جدول‌های فوق با خروجی واقعی اجرای مدل روی داده‌های مطالعه موردی جایگزین شوند.

۵. نتیجه‌گیری و پیشنهادات

۵-۱. نتیجه‌گیری

پژوهش حاضر با هدف ارائه یک مدل هوشمند مبتنی بر سیستم اطلاعات جغرافیایی و الگوریتم‌های یادگیری ماشین برای شناسایی و پیش‌بینی نقاط پرخطر شهری انجام شد. مسئله اصلی پژوهش از این واقعیت ناشی می‌شود که الگوی توزیع خطر در محیط شهری تابع یک عامل منفرد نیست، بلکه حاصل تعامل مجموعه‌ای از عوامل مکانی، کالبدی، جمعیتی، حمل‌ونقلی و محیطی است. ازاین‌رو، استفاده از روش‌های سنتی که صرفاً بر شناسایی نقاطی با سابقه بالای رخداد تمرکز دارند، نمی‌تواند به‌طور کامل قابلیت پیش‌بینی خطر در محدوده‌های مختلف شهری را فراهم کند.

در چارچوب پیشنهادی، GIS به‌عنوان بستر اصلی مدیریت و تحلیل داده‌های مکانی و الگوریتم‌های یادگیری ماشین به‌عنوان ابزار استخراج الگو و پیش‌بینی مورد استفاده قرار گرفتند. این تلفیق امکان آن را فراهم می‌کند که ابتدا روابط و خوشه‌های مکانی شناسایی شده و سپس با استفاده از متغیرهای استخراج‌شده از محیط پیرامون، احتمال قرارگیری هر واحد مکانی در طبقات مختلف خطر پیش‌بینی شود. چنین رویکردی با مطالعات پیشین همخوان است؛ برای نمونه، آتوموآ و همکاران نشان دادند که ترکیب تحلیل فضایی و Random Forest می‌تواند برای شناسایی و پیش‌بینی نقاط پرخطر مورد استفاده قرار گیرد و مدل آن‌ها در مرحله اعتبارسنجی به دقت ۷۶/۷ درصد دست یافت (۱).

یکی از یافته‌های مهم پژوهش حاضر، اهمیت بالای ساختار فضایی و ویژگی‌های محیط پیرامون نقاط مورد مطالعه است. یک نقطه شهری را نمی‌توان مستقل از شبکه معابر، کاربری‌های اطراف، تراکم جمعیت، مراکز فعالیت و سابقه رخدادهای مجاور تحلیل کرد. این موضوع در پژوهش وو و همکاران نیز تأیید شده است. آنان با تلفیق Space-Time Cube و XGBoost نشان دادند که ویژگی‌های نقاط همسایه نقش مهمی در پیش‌بینی نقاط پرخطر دارند و لحاظ کردن روابط مکانی–زمانی می‌تواند دقت پیش‌بینی را نسبت به روش‌های متعارف افزایش دهد (۲).

از منظر الگوریتمی، استفاده از مدل‌هایی مانند Random Forest، SVM و XGBoost این امکان را ایجاد می‌کند که روابط غیرخطی میان عوامل خطر شناسایی شود. بر اساس نتایج نمونه ارائه‌شده در بخش یافته‌ها، XGBoost عملکرد بهتری نسبت به دو مدل دیگر نشان داد. البته این مقادیر در بخش نتایج به‌صورت نمونه ارائه شده‌اند و برای گزارش نهایی مقاله باید با خروجی واقعی اجرای مدل روی داده‌های مطالعه موردی جایگزین شوند. بنابراین، نمی‌توان عملکرد عددی نمونه را بدون اجرای مدل به‌عنوان نتیجه قطعی پژوهش تلقی کرد.

از سوی دیگر، تحلیل اهمیت متغیرها نشان می‌دهد که متغیرهایی نظیر سابقه رخداد، تراکم تقاطع‌ها، شدت تردد، کاربری زمین، تراکم جمعیت و ویژگی‌های شبکه معابر می‌توانند نقش مهمی در افزایش احتمال خطر داشته باشند. این یافته از نظر برنامه‌ریزی شهری اهمیت زیادی دارد، زیرا خروجی مدل تنها یک نقشه خطر نیست؛ بلکه می‌تواند عوامل ایجادکننده خطر را نیز مشخص کند. در نتیجه، مدیر شهری می‌تواند به جای اجرای اقدامات یکسان در کل شهر، مداخلات خود را بر اساس علت خطر در هر محدوده طراحی کند.

یکی دیگر از جنبه‌های مهم مدل پیشنهادی، استفاده از روش‌های تفسیرپذیری مانند SHAP است. مدل‌های پیچیده یادگیری ماشین در صورت استفاده بدون روش‌های تفسیرپذیری می‌توانند به‌صورت «جعبه سیاه» عمل کنند. SHAP امکان بررسی سهم نسبی متغیرها در خروجی مدل را فراهم می‌کند و به تصمیم‌گیرندگان نشان می‌دهد که افزایش یا کاهش یک متغیر چگونه احتمال قرارگیری یک منطقه در طبقه خطر را تغییر می‌دهد. مطالعه وو و همکاران نیز از SHAP برای تفسیر عوامل مؤثر بر نقاط پرخطر و بررسی نقش امکانات شهری، حمل‌ونقل عمومی و کاربری زمین استفاده کرده است (۲).

از منظر نظری، نتایج این پژوهش نشان می‌دهد که آینده تحلیل خطر شهری به سمت تلفیق سه حوزه اصلی یعنی تحلیل مکانی، هوش مصنوعی و تصمیم‌گیری شهری حرکت می‌کند. مرور اکیانیو و همکاران نیز نشان داده است که ترکیب GIS و یادگیری ماشین ظرفیت قابل توجهی برای حل مسائل پیچیده مکانی و توسعه مدل‌های پیش‌بینی دارد (۳).

در عین حال، توسعه چنین مدل‌هایی باید همراه با توجه به کیفیت داده، شفافیت الگوریتم و ملاحظات اخلاقی باشد. پژوهش ماراسینگه و همکاران در زمینه GeoAI شهری بر اهمیت تفسیرپذیری، شفافیت، حریم خصوصی، حکمرانی داده، ایمنی و پاسخگویی در کاربردهای هوش مصنوعی مکانی تأکید می‌کند (۴).

در مجموع، می‌توان نتیجه گرفت که مدل پیشنهادی می‌تواند از یک سیستم صرفاً توصیفی برای نمایش وضعیت موجود به یک سامانه تصمیم‌یار پیش‌بینانه تبدیل شود. این سامانه می‌تواند نقاطی را که احتمال افزایش خطر در آن‌ها وجود دارد، پیش از وقوع رخدادهای بیشتر شناسایی کند و در نتیجه زمینه را برای مدیریت پیشگیرانه فراهم سازد. مهم‌ترین ارزش کاربردی مدل نیز در همین انتقال از «واکنش پس از رخداد» به «پیشگیری مبتنی بر داده» نهفته است.

۲-۵. پیشنهادهای کاربردی

۱. پیشنهاد می‌شود شهرداری‌ها و سازمان‌های مدیریت شهری، بانک اطلاعات مکانی یکپارچه‌ای از رخدادهای خطرناک، شبکه معابر، کاربری زمین، جمعیت، حمل‌ونقل عمومی و زیرساخت‌های شهری ایجاد کنند.

۲. به‌جای استفاده صرف از نقشه‌های ایستا، از مدل‌های یادگیری ماشین برای به‌روزرسانی دوره‌ای نقشه احتمال خطر استفاده شود.

۳. در مناطق دارای احتمال خطر بالا، مطالعات میدانی برای شناسایی علت اصلی خطر انجام شود و مداخلات متناسب با عامل غالب طراحی گردد.

۴. در تحلیل نقاط پرخطر، ویژگی‌های مکانی واحدهای همجوار نیز لحاظ شود؛ زیرا پژوهش‌های جدید نشان داده‌اند که ویژگی‌های نقاط پیرامونی می‌توانند نقش قابل توجهی در پیش‌بینی خطر داشته باشند (۲).

۵. برای انتخاب مدل نهایی، صرفاً Accuracy ملاک قرار نگیرد و شاخص‌هایی مانند Precision، Recall، F1-Score و AUC نیز مورد توجه قرار گیرند.

۶. استفاده از روش‌های Explainable AI مانند SHAP برای افزایش قابلیت اعتماد مدیران شهری به خروجی مدل توصیه می‌شود.

۷. مدل پیشنهادی در دوره‌های زمانی مختلف مجدداً آموزش داده شود تا تغییرات الگوی جمعیت، کاربری زمین، شبکه حمل‌ونقل و رخدادهای شهری در پیش‌بینی لحاظ شود.

۸. برای افزایش قابلیت تعمیم مدل، داده‌های چندمنبعی مانند تصاویر ماهواره‌ای، داده‌های ترافیکی، داده‌های جمعیتی، اطلاعات حسگرهای شهری و داده‌های مکانی رخدادها با یکدیگر تلفیق شوند.

۵-۳. پیشنهاد برای پژوهش‌های آینده

در پژوهش‌های آینده می‌توان مدل پیشنهادی را با الگوریتم‌های پیشرفته‌تر یادگیری عمیق و مدل‌های مکانی – زمانی توسعه داد. به‌طور خاص، استفاده از شبکه‌های عصبی گرافی می‌تواند برای مدل‌سازی ارتباط میان شبکه معابر و واحدهای فضایی مورد بررسی قرار گیرد. همچنین می‌توان از LSTM یا مدل‌های ترکیبی برای پیش‌بینی تغییرات زمانی خطر استفاده کرد.

پیشنهاد دیگر، توسعه مدل به‌صورت چندمخاطره‌ای است؛ به‌گونه‌ای که به جای تمرکز صرف بر یک نوع رخداد، خطرهای مختلف شهری نظیر تصادفات، سیلاب، آتش‌سوزی و سایر رخدادهای اضطراری در یک چارچوب مشترک تحلیل شوند.

همچنین پیشنهاد می‌شود در مطالعات آینده اثر مقیاس مکانی مورد بررسی قرار گیرد. انتخاب اندازه نامناسب واحدهای مکانی می‌تواند بر نتایج مدل اثر بگذارد؛ بنابراین اجرای مدل در چند مقیاس و مقایسه نتایج می‌تواند به تعیین مقیاس بهینه تحلیل کمک کند.

۵-۴. محدودیت‌های پژوهش

مهم‌ترین محدودیت چنین مطالعاتی، وابستگی مدل به کیفیت و کامل بودن داده‌های مکانی است. خطا در مختصات رخدادها، ثبت ناقص حوادث، تفاوت در مقیاس لایه‌های مکانی و عدم تعادل بین نمونه‌های پرخطر و غیرپرخطر می‌تواند عملکرد مدل را تحت تأثیر قرار دهد. همچنین، نتایج یک مدل یادگیری ماشین الزاماً در تمام شهرها قابل تعمیم نیست و لازم است مدل در هر شهر با داده‌های همان محدوده آموزش و اعتبارسنجی شود.

محدودیت دیگر مربوط به تفسیر رابطه میان متغیرها و خطر است. اهمیت بالای یک متغیر در مدل لزوماً به معنای وجود رابطه علّی مستقیم نیست؛ بنابراین نتایج مدل باید در کنار دانش تخصصی حوزه شهری و مطالعات میدانی تفسیر شوند.

References

Abdel-Aty, M., Siddiqui, C., Huang, H., & Wang, X. (2013). Geographical unit based analysis in the context of transportation safety planning. *Transportation Research Part A: Policy and Practice, 49*, 62–75.
https://doi.org/10.1016/j.tra.2013.01.030.

Aguero-Valverde, J., & Jovanis, P. P. (2006). Spatial analysis of fatal and injury crashes in Pennsylvania. *Accident Analysis & Prevention, 38*(4), 618–625.
https://doi.org/10.1016/j.aap.2005.12.006.

Atumoa, E. A., Fang, T., & Jiang, X. (2022). Spatial statistics and random forest approaches for traffic crash hot spot identification and prediction. *International Journal of Injury Control and Safety Promotion, 29*(2), 207–216.
https://doi.org/10.1080/17457300.2021.1983844.

Ekeanyanwu, C. V., Obisakin, I. F., Aduwenye, P., & Dede-Bamfo, N. (2022). Merging GIS and machine learning techniques: A paper review. *Journal of Geoscience and Environment Protection, 10*, 61–83.
https://doi.org/10.4236/gep.2022.109004.

Marasinghe, R., Yigitcanlar, T., Mayere, S., Washington, T., & Limb, M. (2024). Towards responsible urban geospatial AI: Insights from the white and grey literatures. *Journal of Geovisualization and Spatial Analysis, 8*, 24.
https://doi.org/10.1007/s41651-024-00184-2.

Marasinghe, R., Yigitcanlar, T., Mayere, S., Washington, T., & Limb, M. (2024). Towards responsible urban geospatial AI: Insights from the white and grey literatures. *Journal of Geovisualization and Spatial Analysis, 8*, 24.
https://doi.org/10.1007/s41651-024-00184-2.

Wu, P., Chen, T., Diew Wong, Y., Meng, X., Wang, X., & Liu, W. (2023). Exploring key spatio-temporal features of crash risk hot spots on urban road network: A machine learning approach. *Transportation Research Part A: Policy and Practice, 173*, 103717.
https://doi.org/10.1016/j.tra.2023.103717.

Wu, P., Chen, T., Wong, Y. D., Meng, X., Wang, X., & Liu, W. (2023). Exploring key spatio-

temporal features of crash risk hot spots on urban road network: A machine learning approach. *Transportation Research Part A: Policy and Practice, 173*, 103717. https://doi.org/10.1016/j.tra.2023.103717.