

بررسی کاربرد ترکیبی GIS و یادگیری ماشین در تحلیل و پیش‌بینی الگوهای توسعه شهری

مقدمه

توسعه شهری یک فرایند صرفاً کالبدی نیست؛ بلکه حاصل برهم‌کنش مجموعه‌ای از عوامل اقتصادی، اجتماعی، جمعیتی، محیط‌زیستی، زیرساختی، شبکه حمل‌ونقل، سیاست‌های زمین و تصمیمات مدیریتی است. همین چندعاملی بودن باعث شده است که تحلیل مسیر رشد شهرها و پیش‌بینی جهت توسعه آینده، یکی از مسائل پیچیده در برنامه‌ریزی شهری و منطقه‌ای باشد.

در گذشته، بخش قابل‌توجهی از تحلیل‌های مربوط به توسعه شهری بر نقشه‌های سنتی، آمارهای جمعیتی، تحلیل‌های آماری و مدل‌های مبتنی بر قواعد از پیش تعیین‌شده استوار بود. با گسترش سامانه‌های اطلاعات جغرافیایی (GIS)، امکان ترکیب داده‌های مکانی مختلف و تحلیل هم‌زمان آن‌ها فراهم شد. به برنامه‌ریزان اجازه می‌دهد داده‌هایی مانند کاربری زمین، شبکه معابر، فاصله از مراکز شهری، شیب، زیرساخت‌ها، تراکم جمعیت و محدوده‌های ساخته‌شده را در یک محیط مکانی یکپارچه کنند.

با این حال، GIS به‌تنهایی یک روش پیش‌بینی‌کننده نیست. در اینجا یادگیری ماشین اهمیت پیدا می‌کند. الگوریتم‌های یادگیری ماشین می‌توانند از الگوهای موجود در داده‌های تاریخی یاد بگیرند و رابطه میان متغیرهای مکانی و احتمال تغییرات آتی را استخراج کنند. ترکیب GIS و یادگیری ماشین در واقع یک گام به سمت **هوش مکانی** یا **GeoAI** محسوب می‌شود؛ حوزه‌ای که در سال‌های اخیر برای تحلیل داده‌های مکانی، سنجش‌ازدور و فرایندهای پیچیده شهری اهمیت بیشتری پیدا کرده است.

هدف این مقاله بررسی این است که GIS و یادگیری ماشین چگونه می‌توانند در کنار یکدیگر برای تشخیص الگوهای توسعه شهری، شناسایی عوامل مؤثر بر رشد، طبقه‌بندی تغییرات کاربری زمین و پیش‌بینی مناطق مستعد توسعه آینده به کار گرفته شوند. همچنین محدودیت‌ها، روش‌های ارزیابی و یک چارچوب پیشنهادی برای طراحی پژوهش دانشگاهی در این زمینه بررسی می‌شود.

GIS و یادگیری ماشین؛ چرا ترکیب این دو اهمیت دارد؟

GIS را می‌توان بستر مدیریت، سازمان‌دهی، تحلیل و نمایش داده‌های دارای مؤلفه مکانی دانست. برای مثال، در یک مطالعه توسعه شهری می‌توان لایه‌های زیر را در GIS آماده کرد:

- محدوده ساخته شده در دوره های زمانی مختلف

- کاربری اراضی

- شبکه معابر

- فاصله از راه های اصلی

- فاصله از مرکز شهر

- فاصله از مراکز خدماتی

- تراکم جمعیت

- تراکم ساختمانی

- شیب و ارتفاع

- رودخانه ها و پهنه های طبیعی

- محدوده های ممنوعه یا حفاظت شده

- قیمت زمین یا شاخص های اقتصادی

- ایستگاه های حمل و نقل عمومی

اما پرسش اصلی این است که ترکیب این متغیرها چگونه به تغییر کاربری یا توسعه شهری منجر می شود؟

اگر رابطه میان این عوامل ساده و خطی بود، استفاده از مدل های آماری کلاسیک ممکن بود کافی باشد. اما در بسیاری از فرایندهای شهری، رابطه ها غیر خطی، چندمتغیره و وابسته به موقعیت مکانی هستند. برای مثال، فاصله از یک جاده اصلی ممکن است تا یک محدوده خاص احتمال توسعه را افزایش دهد، اما پس از آن عوامل دیگری مانند شیب، محدودیت قانونی یا فاصله از مرکز شهری اهمیت بیشتری پیدا کنند.

یادگیری ماشین برای کشف چنین روابطی مناسب است. در این رویکرد، GIS عمدتاً نقش تولید و مدیریت داده های مکانی را ایفا می کند و الگوریتم یادگیری ماشین نقش یادگیری رابطه میان متغیرهای ورودی و پدیده مورد مطالعه را بر عهده می گیرد.

این ترکیب را می‌توان به صورت ساده چنین نمایش داد:

داده‌های مکانی → آماده‌سازی در GIS → استخراج متغیرهای مؤثر → آموزش مدل یادگیری ماشین → تولید احتمال یا پیش‌بینی مکانی → بازگرداندن نتایج به GIS تحلیل و تفسیر شهری

این چارچوب، GIS را از یک ابزار صرفاً ترسیمی به بخشی از یک فرایند تحلیلی و پیش‌بینی‌کننده تبدیل می‌کند.

توسعه شهری به عنوان یک مسئله مکانی - زمانی

یکی از نکات مهم در پژوهش‌های مربوط به توسعه شهری این است که رشد شهر را نباید صرفاً به عنوان یک مسئله مکانی در نظر گرفت. شهر در طول زمان تغییر می‌کند و بنابراین داده‌های تاریخی اهمیت زیادی دارند. برای مثال، فرض کنید محدوده ساخته شده یک شهر در سال‌های ۱۳۸۵، ۱۳۹۰، ۱۳۹۵ و ۱۴۰۰ در دسترس باشد. با مقایسه این نقشه‌ها می‌توان مشخص کرد:

۱. کدام زمین‌ها از غیرشهری به شهری تبدیل شده‌اند؟
 ۲. توسعه بیشتر در چه جهتی اتفاق افتاده است؟
 ۳. فاصله مناطق جدید از معابر اصلی چقدر بوده است؟
 ۴. آیا توسعه به شکل پیوسته و متراکم رخ داده یا پراکنده بوده است؟
 ۵. کدام عوامل بیشترین ارتباط را با توسعه داشته‌اند؟
- در چنین حالتی، مسئله یادگیری ماشین می‌تواند به شکل یک مسئله طبقه‌بندی یا احتمال‌سنجی تعریف شود.

برای نمونه:

متغیر هدف:

آیا یک سلول یا قطعه زمین در دوره آینده به محدوده ساخته شده تبدیل خواهد شد یا خیر؟

متغیرهای ورودی:

- فاصله تا شبکه معابر
- فاصله تا محدوده ساخته شده موجود
- فاصله تا مرکز شهر
- شیب
- کاربری فعلی
- تراکم اطراف
- فاصله تا مراکز خدماتی
- فاصله تا ایستگاه حمل و نقل
- ویژگی های جمعیتی
- متغیرهای اقتصادی و اجتماعی

در این حالت، مدل یادگیری ماشین می تواند برای هر موقعیت مکانی، احتمال توسعه را برآورد کند.

نقش سنجش از دور در کنار GIS و یادگیری ماشین

یکی از مهم ترین منابع داده برای مطالعات توسعه شهری، تصاویر ماهواره ای و داده های سنجش از دور هستند. تصاویر چندزمانه این امکان را ایجاد می کنند که تغییرات پوشش و کاربری زمین در طول زمان استخراج شوند.

در یک چارچوب پژوهشی معمول می توان از تصاویر ماهواره ای برای تولید نقشه های کاربری یا پوشش زمین استفاده کرد و سپس این نقشه ها را به GIS منتقل کرد.

یادگیری ماشین در این مرحله نیز کاربرد دارد. الگوریتم هایی مانند **Support Random Forest** و **Vector Machine** شبکه های عصبی برای طبقه بندی داده های سنجش از دور و تشخیص الگوهای کاربری زمین مورد استفاده قرار گرفته اند. مرورهای پژوهشی در حوزه برنامه ریزی کاربری زمین نیز نشان

می‌دهند که RF، SVM و CNN از روش‌های مهم در تحلیل داده‌های سنجش‌ازدور و مطالعات الگوی فضایی هستند.

بنابراین یک زنجیره کامل پژوهشی می‌تواند چنین باشد:

تصاویر ماهواره‌ای → طبقه‌بندی کاربری زمین → استخراج تغییرات → ایجاد لایه‌های GIS → استخراج عوامل مؤثر → آموزش مدل → ML پیش‌بینی توسعه آینده

این ساختار به‌خصوص برای شهرهایی که اطلاعات تاریخی دقیق از کاربری زمین در آن‌ها محدود است، اهمیت زیادی دارد.

الگوریتم‌های یادگیری ماشین مورد استفاده در تحلیل توسعه شهری

انتخاب الگوریتم نباید صرفاً بر اساس محبوبیت آن انجام شود. نوع داده، حجم نمونه، ساختار فضایی مسئله، هدف پژوهش و قابلیت تفسیر مدل باید در انتخاب الگوریتم مؤثر باشند.

۱. رگرسیون لجستیک

رگرسیون لجستیک یکی از روش‌های ساده و قابل تفسیر برای پیش‌بینی احتمال وقوع یک رویداد است. در مسئله توسعه شهری می‌توان از آن برای برآورد احتمال تبدیل یک سلول از «غیرساخته‌شده» به «ساخته‌شده» استفاده کرد.

مزیت مهم این روش، قابلیت تفسیر رابطه بین متغیرهاست. بنابراین برای پژوهش‌هایی که علاوه بر پیش‌بینی، به دنبال توضیح عوامل مؤثر هستند، نقطه شروع مناسبی محسوب می‌شود.

با این حال، اگر روابط میان متغیرها بسیار پیچیده و غیرخطی باشند، مدل‌های پیشرفته‌تر می‌توانند عملکرد متفاوتی داشته باشند.

۲. درخت تصمیم

درخت تصمیم مسئله را از طریق مجموعه‌ای از قواعد تصمیم‌گیری تقسیم می‌کند.

برای مثال، مدل ممکن است به‌صورت مفهومی یاد بگیرد:

اگر فاصله از جاده کم باشد و زمین خارج از محدوده محدودیت قرار داشته باشد، احتمال توسعه بیشتر است. مزیت درخت تصمیم این است که منطق آن نسبتاً قابل مشاهده است. در مطالعات برنامه‌ریزی شهری که تفسیرپذیری اهمیت زیادی دارد، این ویژگی ارزشمند است.

۳. Random Forest

Random Forest مجموعه‌ای از درخت‌های تصمیم است که به صورت ترکیبی عمل می‌کنند. این الگوریتم برای داده‌های چندمتغیره و روابط غیرخطی بسیار مورد استفاده قرار گرفته است.

در مطالعات پیش‌بینی توسعه شهری، می‌توان از Random Forest برای برآورد احتمال توسعه بر اساس مجموعه‌ای از عوامل مکانی استفاده کرد.

یکی از مزیت‌های مهم آن این است که می‌تواند اهمیت نسبی متغیرها را نیز بررسی کند. در نتیجه پژوهشگر می‌تواند علاوه بر سؤال «کجا احتمال توسعه وجود دارد؟»، تا حدی به سؤال «کدام عوامل در پیش‌بینی نقش بیشتری داشته‌اند؟» نیز پاسخ دهد.

مرورهای جدیدتر درباره یادگیری ماشین در پیش‌بینی کاربری زمین نیز استفاده گسترده از Random Forest، در کنار مدل‌هایی مانند SVM و شبکه‌های عصبی، را نشان می‌دهند.

۴. Support Vector Machine

SVM به‌ویژه در مسائل طبقه‌بندی داده‌های پیچیده کاربرد دارد.

در مطالعات شهری می‌توان از آن برای طبقه‌بندی کاربری زمین یا تفکیک مناطق دارای پتانسیل توسعه استفاده کرد.

SVM در شرایطی که تعداد ویژگی‌ها نسبتاً زیاد و حجم نمونه محدودتر باشد، می‌تواند گزینه قابل بررسی باشد؛ البته انتخاب هسته، تنظیم پارامترها و آماده‌سازی داده بر عملکرد آن اثر قابل توجهی دارد.

۵. شبکه‌های عصبی و یادگیری عمیق

با افزایش حجم داده‌های مکانی و تصویری، مدل‌های یادگیری عمیق نیز در مطالعات شهری اهمیت بیشتری پیدا کرده‌اند.

به‌ویژه **Convolutional Neural Networks** یا **CNN** برای داده‌های تصویری و ساختارهای مکانی مناسب هستند. CNN می‌تواند الگوهای محلی را از تصاویر و داده‌های شبکه‌ای استخراج کند.

در مسائل پیچیده‌تر، مدل‌های عمیق می‌توانند برای استخراج ویژگی‌های فضایی، تحلیل تصاویر ماهواره‌ای یا ترکیب داده‌های چندمنبعی مورد استفاده قرار گیرند.

با این حال، افزایش پیچیدگی مدل الزاماً به معنای بهتر شدن پژوهش نیست. مدل‌های عمیق معمولاً به داده و توان محاسباتی بیشتری نیاز دارند و تفسیر خروجی آن‌ها نیز دشوارتر است.

مدل‌های ترکیبی؛ پیوند یادگیری ماشین با Cellular Automata

یکی از جذاب‌ترین کاربردهای ترکیبی در مدل‌سازی توسعه شهری، اتصال یادگیری ماشین با اتوماتای سلولی یا **Cellular Automata (CA)** است.

در مدل CA، فضای شهری به سلول‌هایی تقسیم می‌شود و وضعیت هر سلول می‌تواند در طول زمان تغییر کند. تغییر وضعیت هر سلول تحت تأثیر وضعیت خودش، همسایگی آن و مجموعه‌ای از قواعد انتقال قرار می‌گیرد.

CA از سال‌ها پیش در شبیه‌سازی رشد شهری کاربرد داشته است و سازگاری بالایی با محیط‌های GIS دارد.

مشکل اصلی این است که تعیین قواعد انتقال مناسب می‌تواند دشوار باشد.

اینجاست که یادگیری ماشین وارد می‌شود.

به جای تعیین دستی قواعد، می‌توان از داده‌های تاریخی استفاده کرد تا مدل ML احتمال انتقال هر سلول را یاد بگیرد. سپس این احتمال به‌عنوان بخشی از قواعد انتقال در مدل CA استفاده شود.

در نتیجه، یک مدل ترکیبی می‌تواند ساختاری مشابه زیر داشته باشد:

GIS + داده‌های تاریخی = Machine Learning + Cellular Automata + مدل پیش‌بینی
رشد شهری

این رویکرد نسبت به استفاده مستقل از هر کدام از روش‌ها، امکان ترکیب یادگیری از داده با شبیه‌سازی صریح فضایی را فراهم می‌کند.

مطالعات مروری درباره مدل‌های رشد و تراکم شهری نیز بر اهمیت مدل‌های ترکیبی، از جمله ترکیب CA با روش‌های آماری، شبکه عصبی و رگرسیون، تأکید کرده‌اند.

مراحل اجرای یک پژوهش GIS - Machine Learning برای توسعه شهری

برای دانشجوی مهندسی شهرسازی، داشتن یک روند روش‌شناختی مشخص اهمیت زیادی دارد.

مرحله اول: تعریف دقیق مسئله

ابتدا باید مشخص شود که دقیقاً چه چیزی قرار است پیش‌بینی شود.

برای مثال:

- توسعه محدوده ساخته‌شده
- تغییر کاربری کشاورزی به شهری
- گسترش پراکنده شهر
- توسعه پیرامون محورهای حمل‌ونقل
- احتمال توسعه در قطعات خالی
- تغییر الگوی تراکم شهری

عبارت «پیش‌بینی توسعه شهری» به‌تنهایی مسئله پژوهشی کاملی نیست.

مرحله دوم: تعیین واحد تحلیل

واحد تحلیل می‌تواند یکی از موارد زیر باشد:

- سلول رستری

- قطعه زمین

- بلوک شهری

- ناحیه آماری

- منطقه شهری

- واحدهای شبکه‌ای منظم

انتخاب واحد تحلیل بر نوع مدل و نتایج تأثیر می‌گذارد.

برای مثال، اگر مدل بر اساس رستر اجرا شود، مسئله به‌شکل مجموعه‌ای از سلول‌های مکانی تعریف خواهد شد؛ درحالی‌که تحلیل در سطح قطعه زمین می‌تواند ویژگی‌های هندسی و مالکیتی را نیز وارد مدل کند.

مرحله سوم: آماده‌سازی داده‌های تاریخی

برای آموزش مدل، معمولاً به داده‌های مربوط به چند مقطع زمانی نیاز است.

به‌عنوان نمونه:

زمان ۱ → زمان ۲ → زمان ۳

سپس تغییرات بین زمان‌های مختلف استخراج می‌شوند.

نکته بسیار مهم این است که داده آموزش و داده آزمون نباید به‌گونه‌ای تقسیم شوند که اطلاعات یکسان یا بسیار مشابه از یک مکان در هر دو بخش قرار گیرد؛ زیرا این مسئله می‌تواند باعث **Data Leakage** و برآورد بیش از حد عملکرد مدل شود.

مرحله چهارم: ایجاد متغیرهای توضیحی

در این مرحله، GIS نقش کلیدی دارد.

برای هر سلول یا قطعه زمین، می‌توان متغیرهایی مانند موارد زیر تولید کرد:

نمونه متغیر	گروه متغیر
فاصله تا معبر اصلی	دسترسی
فاصله تا ایستگاه مترو یا اتوبوس	حمل و نقل
فاصله تا مرکز شهر	مرکزیت
تراکم ساخت و ساز اطراف	کالبدی
شیب، ارتفاع	محیطی
نوع کاربری فعلی	کاربری
تراکم جمعیت	اجتماعی
فاصله تا مراکز خدمات	خدماتی
فاصله از پهنه‌های ممنوعه	محدودیت
درصد ساخت و ساز در سلول‌های اطراف	همسایگی

این مرحله در بسیاری از پروژه‌ها حتی از انتخاب الگوریتم مهم‌تر است؛ زیرا یک مدل پیچیده با داده‌های نامناسب نمی‌تواند یک مسئله بدتعریف‌شده را اصلاح کند.

آموزش مدل و تولید نقشه احتمال توسعه

پس از آماده‌سازی داده، نمونه‌ها به دو بخش یا چند بخش تقسیم می‌شوند:

داده آموزش **Training** →

داده اعتبارسنجی **Validation** →

داده آزمون **Test** →

مدل با داده‌های آموزش یاد می‌گیرد و سپس روی داده‌ای که در فرایند آموزش استفاده نشده ارزیابی می‌شود.

خروجی مدل الزاماً نباید یک نقشه دوحالتی «توسعه/عدم توسعه» باشد. در بسیاری از کاربردها، خروجی بهتر است به شکل یک سطح احتمال یا پتانسیل توسعه تولید شود.

برای مثال:

- احتمال پایین
- احتمال متوسط
- احتمال بالا

یا یک سطح پیوسته بین صفر و یک.

این سطح سپس می‌تواند در GIS نمایش داده شود و همراه با محدودیت‌های برنامه‌ریزی و سناریوهای مختلف تحلیل شود.

ارزیابی عملکرد مدل

یکی از اشتباهات رایج در پژوهش‌های یادگیری ماشین، تمرکز صرف بر Accuracy است.

فرض کنید تنها ۱۰ درصد منطقه مورد مطالعه در دوره بعدی توسعه پیدا کند. مدلی که تقریباً همه مناطق را «بدون توسعه» اعلام کند ممکن است Accuracy بالایی به دست آورد، اما برای برنامه‌ریزی شهری مدل مفیدی نباشد.

بنابراین باید از معیارهای متناسب با مسئله استفاده کرد، از جمله:

- Accuracy
- Precision
- Recall
- F₁-score
- ROC-AUC
- PR-AUC
- Specificity

در کنار این معیارها، در مسائل مکانی باید اعتبارسنجی فضایی نیز جدی گرفته شود.

اگر نقاط آموزش و آزمون از نظر مکانی بسیار نزدیک باشند، ممکن است مدل ظاهراً عملکرد بالایی داشته باشد، در حالی که در یک منطقه جغرافیایی جدید عملکرد آن افت کند.

بنابراین برای پژوهش‌های دانشگاهی بهتر است علاوه بر اعتبارسنجی معمول، امکان **Spatial Cross-Validation** یا اعتبارسنجی مبتنی بر تفکیک فضایی نیز بررسی شود.

تحلیل الگوی توسعه شهری پس از پیش‌بینی

نقشه خروجی مدل پایان کار نیست.

یک پژوهش شهرسازی باید از خروجی الگوریتم به تحلیل فضایی و برنامه‌ریزی برسد.

برای مثال، اگر مدل مناطق مستعد توسعه را شناسایی کند، می‌توان پرسید:

- آیا این مناطق با شبکه حمل‌ونقل هماهنگ هستند؟
- آیا در محدوده‌های دارای ارزش اکولوژیک قرار گرفته‌اند؟
- آیا توسعه در آن‌ها باعث گسترش افقی شهر می‌شود؟
- آیا ظرفیت زیرساختی برای توسعه وجود دارد؟
- آیا توسعه پیشنهادی با طرح‌های بالادست سازگار است؟
- آیا توسعه باعث افزایش پراکندگی شهری می‌شود؟
- آیا مناطق مستعد توسعه از خدمات شهری فاصله زیادی دارند؟

این مرحله نشان می‌دهد که پیش‌بینی ماشین یادگیری، جایگزین تصمیم‌گیری شهرسازی نیست.

مدل می‌تواند شواهد و سناریو تولید کند، اما تفسیر آن باید با دانش برنامه‌ریزی، ضوابط، سیاست‌های زمین، ملاحظات اجتماعی و محدودیت‌های محیطی ترکیب شود.

کاربرد GIS و یادگیری ماشین در تحلیل پراکنش شهری

یکی از کاربردهای مهم این ترکیب، بررسی **Urban Sprawl** یا پراکنش شهری است.

در پراکنش شهری، تنها میزان رشد سطح ساخته شده مهم نیست؛ بلکه شکل، پیوستگی و جهت رشد نیز اهمیت دارد.

با استفاده از GIS می توان شاخص های مختلفی برای سنجش الگوی توسعه تولید کرد و سپس آن ها را در مدل های یادگیری ماشین وارد کرد.

برای مثال، می توان بررسی کرد که آیا توسعه جدید:

- به هسته شهری متصل است؛
- در امتداد شبکه حمل و نقل ایجاد شده است؛
- به صورت لکه ای و پراکنده شکل گرفته است؛
- یا در قالب توسعه پیوسته و فشرده رخ داده است.

پژوهش های مربوط به مورفولوژی شهری نیز بر نقش GIS در تحلیل پیکربندی فضایی، اتصال پذیری، تغییرات زمانی و ویژگی های کالبدی تأکید دارند.

بنابراین ترکیب معیارهای مورفولوژیک با مدل های ML می تواند تصویری غنی تر از صرفاً «میزان رشد شهر» ارائه دهد.

مزایا و محدودیت های رویکرد ترکیبی

مزایا

۱. استفاده هم زمان از داده های متنوع

GIS می تواند داده های مکانی با منابع مختلف را یکپارچه کند و یادگیری ماشین از همین مجموعه چندمتغیره برای آموزش استفاده کند.

۲. شناسایی روابط غیر خطی

برخی الگوریتم ها می توانند روابط پیچیده میان عوامل توسعه را بدون نیاز به تعریف تمام روابط به صورت دستی یاد بگیرند.

۳. امکان تولید نقشه های پیش بینی

خروجی مدل را می‌توان دوباره به محیط GIS منتقل کرد و به‌صورت نقشه احتمال یا پتانسیل توسعه نمایش داد.

۴. امکان مقایسه سناریوها

می‌توان برخی متغیرها را تغییر داد و اثر سناریوهای مختلف را بر الگوی توسعه بررسی کرد.

۵. افزایش ظرفیت تحلیل داده‌های بزرگ

با افزایش داده‌های ماهواره‌ای، شبکه‌های حمل‌ونقل، داده‌های شهری و منابع مکانی، روش‌های خودکار اهمیت بیشتری پیدا می‌کنند.

محدودیت‌ها

۱. کیفیت داده

مدل یادگیری ماشین نمی‌تواند به‌طور جادویی داده‌های ناقص، نادرست یا ناسازگار را اصلاح کند.

۲. سوگیری داده

اگر داده‌های تاریخی منعکس‌کننده سیاست‌ها یا نابرابری‌های خاصی باشند، مدل ممکن است همان الگوها را بازتولید کند.

۳. مسئله انتقال‌پذیری

مدلی که در یک شهر عملکرد خوبی دارد، لزوماً در شهر دیگری با ساختار اقتصادی، کالبدی و مدیریتی متفاوت همان عملکرد را ندارد.

۴. تفسیرپذیری

هرچه مدل پیچیده‌تر شود، توضیح اینکه چرا یک موقعیت به‌عنوان منطقه مستعد توسعه شناخته شده است دشوارتر می‌شود.

۵. عدم قطعیت

پیش‌بینی توسعه شهری ذاتاً با عدم قطعیت همراه است. تغییر سیاست‌های زمین، پروژه‌های زیرساختی، شوک‌های اقتصادی و عوامل اجتماعی می‌توانند مسیر توسعه را تغییر دهند.

مرورهای جدید در حوزه مدل سازی شهری با AI نیز بر محدودیت هایی مانند توجه ناکافی به عوامل اجتماعی - اقتصادی، تفسیرپذیری، انتقال پذیری و تبدیل «دقت پیش بینی» به تصمیم پشتیبانی شده تأکید دارند.

یک چارچوب پیشنهادی برای پایان نامه مهندسی شهرسازی

برای دانشجوی کارشناسی ارشد یا دکتری، می توان یک پژوهش را با چارچوب زیر طراحی کرد:

گام اول: تعریف محدوده مطالعه

انتخاب یک شهر یا ناحیه شهری و تعیین دوره زمانی مطالعه.

گام دوم: تهیه داده

جمع آوری:

- تصاویر ماهواره ای
- نقشه کاربری زمین
- شبکه معابر
- اطلاعات جمعیتی
- مراکز خدماتی
- توپوگرافی
- محدوده های محدودیت توسعه
- داده های تاریخی توسعه شهری

گام سوم: استخراج تغییرات

با استفاده از داده های چندزمانه، مناطق توسعه یافته در دوره تاریخی مشخص شوند.

گام چهارم: تولید ویژگی های مکانی

در GIS برای هر واحد تحلیل، متغیرهای مؤثر بر توسعه محاسبه شود.

گام پنجم: ساخت مدل

حداقل دو یا چند الگوریتم مانند:

Logistic Regression + Random Forest + SVM

یا در صورت وجود داده کافی:

Random Forest + CNN

با یکدیگر مقایسه شوند.

گام ششم: ارزیابی

مدل‌ها با معیارهای مناسب ارزیابی و از نظر دقت، تعمیم‌پذیری و تفسیرپذیری مقایسه شوند.

گام هفتم: تولید نقشه پیش‌بینی

بهترین مدل از نظر معیارهای از پیش تعیین‌شده برای تولید نقشه پتانسیل توسعه استفاده شود.

گام هشتم: تحلیل شهرسازی

نقشه پیش‌بینی با:

- طرح‌های توسعه
- شبکه حمل‌ونقل
- خدمات شهری
- پهنه‌های محیط‌زیستی
- مناطق پرخطر
- ظرفیت زیرساختی

مقایسه شود.

در نهایت، هدف پایان‌نامه نباید صرفاً «ساخت یک مدل با Accuracy بالا» باشد؛ بلکه باید مشخص شود مدل چگونه می‌تواند به شناخت فرایند توسعه و پشتیبانی از تحلیل و تصمیم‌سازی فضایی کمک کند.

نقش مدل‌های سناریومحور در پیش‌بینی آینده شهر

یکی از ضعف‌های رایج در برخی مطالعات، برابر دانستن پیش‌بینی آماری با آینده واقعی شهر است.

در برنامه‌ریزی شهری بهتر است به جای طرح این سؤال که:

«شهر دقیقاً در آینده چگونه توسعه خواهد یافت؟»

پرسیده شود:

«با توجه به الگوهای گذشته و تحت مجموعه‌ای از فرضیات مشخص، کدام مناطق پتانسیل بیشتری برای توسعه دارند؟»

این تغییر نگاه بسیار مهم است.

مدل می‌تواند سناریوهای مختلفی تولید کند. برای مثال:

سناریوی توسعه متراکم:

اولویت با توسعه درون‌بافت و زمین‌های خالی داخل محدوده شهری است.

سناریوی توسعه پیرامونی:

فرض می‌شود فشار توسعه به سمت حاشیه شهر افزایش پیدا کند.

سناریوی حمل‌ونقل محور:

دسترسی به شبکه حمل‌ونقل عمومی وزن بیشتری در مدل پیدا می‌کند.

سناریوی حفاظت محیط‌زیست:

مناطق حساس محیطی از فرایند توسعه حذف یا محدود می‌شوند.

در اینجا GIS محیط مناسبی برای مقایسه مکانی سناریوها فراهم می‌کند و ML می‌تواند برای برآورد روابط و پتانسیل توسعه به کار رود.

مطالعات مرور ادبیات در حوزه شبیه‌سازی سناریوهای توسعه شهری نیز نشان داده‌اند که مدل‌های سناریویی برای بررسی مسیرهای مختلف تغییر کاربری زمین و پشتیبانی از برنامه‌ریزی پایدار مورد استفاده قرار گرفته‌اند.

ابزارهای قابل استفاده

یک پروژه دانشگاهی در این زمینه الزاماً به یک نرم‌افزار خاص محدود نیست.

در بخش GIS می‌توان از نرم‌افزارهایی مانند **ArcGIS Pro** یا **QGIS** استفاده کرد.

برای پردازش و مدل‌سازی داده نیز محیط‌هایی مانند **Python** اهمیت زیادی دارند. کتابخانه‌هایی مانند **Scikit-learn** برای مدل‌های کلاسیک و چارچوب‌هایی مانند **TensorFlow** یا **PyTorch** برای یادگیری عمیق قابل استفاده‌اند.

در پروژه‌هایی که حجم داده‌های سنجش‌ازدور بسیار زیاد است، **Google Earth Engine** نیز می‌تواند برای آماده‌سازی و پردازش داده‌های سنجش‌ازدور مفید باشد. این پلتفرم **API** های یادگیری ماشین برای برخی وظایف طبقه‌بندی و رگرسیون داده‌های مشاهده زمین دارد و در صورت نیاز می‌توان داده‌ها را برای آموزش مدل‌های عمیق‌تر به محیط‌های خارجی منتقل کرد.

انتخاب ابزار باید تابع سؤال پژوهش باشد، نه برعکس.

اشتباهات رایج در پژوهش‌های GIS و یادگیری ماشین

استفاده از تعداد زیادی متغیر بدون منطق نظری

افزایش تعداد لایه‌ها الزاماً کیفیت مدل را افزایش نمی‌دهد. هر متغیر باید دلیل نظری یا تجربی قابل دفاعی داشته باشد.

تمرکز صرف بر دقت

مدلی که فقط از نظر یک معیار عددی عملکرد خوبی دارد، لزوماً بهترین مدل برای برنامه‌ریزی شهری نیست.

نادیده گرفتن ساختار فضایی داده

داده‌های شهری مستقل از مکان نیستند. نزدیکی مکانی می‌تواند باعث شباهت نمونه‌ها شود و اعتبارسنجی نادرست، عملکرد مدل را بیش از حد واقعی نشان دهد.

استفاده از داده آینده در آموزش مدل

اگر اطلاعاتی که در زمان پیش‌بینی در دسترس نبوده به‌طور غیرمستقیم وارد متغیرهای آموزشی شود، مدل دچار نشت اطلاعات می‌شود.

یکی دانستن همبستگی و علیت

اگر مدل تشخیص دهد که یک متغیر با توسعه شهری ارتباط دارد، این مسئله به‌تنهایی ثابت نمی‌کند که آن متغیر علت توسعه بوده است.

نادیده گرفتن زمینه شهر

مدل باید در بستر اقتصادی، اجتماعی، کالبدی و نهادی شهر تفسیر شود.

آینده ترکیب GIS و یادگیری ماشین در شهرسازی

مسیر آینده این حوزه احتمالاً به سمت GeoAI، داده‌های چندمنبعی، مدل‌های مکانی - زمانی و یادگیری عمیق حرکت خواهد کرد.

یکی از جهت‌گیری‌های مهم، ترکیب داده‌های مختلف است:

تصاویر ماهواره‌ای + GIS + شبکه حمل‌ونقل + داده‌های جمعیتی + داده‌های اقتصادی + داده‌های محیطی

چنین ترکیبی می‌تواند تصویر جامع‌تری از پویایی شهر ارائه دهد.

از سوی دیگر، روش‌های جدیدتر به سمت مدل‌هایی حرکت می‌کنند که فقط یک تصویر یا یک لایه مکانی را بررسی نمی‌کنند، بلکه الگوهای مکانی و زمانی را به‌صورت هم‌زمان مدل می‌کنند.

موضوع دیگری که اهمیت آن افزایش خواهد یافت، تفسیرپذیری مدل‌ها است. برای برنامه‌ریزی شهری، تنها دانستن اینکه «این منطقه احتمال توسعه دارد» کافی نیست. برنامه‌ریز باید بداند این نتیجه بر اساس چه عواملی حاصل شده است.

در ادبیات جدید GeoAI نیز بر ادغام داده‌های چندمنبعی، مدل‌های مکانی - زمانی، تفسیرپذیری و توسعه چارچوب‌های انسان‌محور تأکید شده است.

جمع‌بندی

ترکیب GIS و یادگیری ماشین را می‌توان یکی از مسیرهای مهم برای توسعه روش‌های نوین تحلیل شهری دانست. GIS توانایی مدیریت و تحلیل داده‌های دارای مؤلفه مکانی را فراهم می‌کند و یادگیری ماشین می‌تواند روابط پیچیده میان متغیرهای مکانی و تغییرات شهری را از داده‌های تاریخی استخراج کند.

در مسئله توسعه شهری، این ترکیب می‌تواند برای شناسایی مناطق مستعد توسعه، تحلیل عوامل مؤثر، طبقه‌بندی کاربری زمین، تشخیص الگوهای پراکنش و تولید سناریوهای احتمالی آینده مورد استفاده قرار گیرد.

با این حال، ارزش یک مدل صرفاً به Accuracy آن وابسته نیست. کیفیت داده، طراحی صحیح نمونه‌ها، جلوگیری از Data Leakage، اعتبارسنجی فضایی، تفسیرپذیری، انتقال‌پذیری و سازگاری نتایج با منطق برنامه‌ریزی شهری همگی باید در نظر گرفته شوند.

برای دانشجوی مهندسی شهرسازی، رویکرد مناسب این نیست که GIS و یادگیری ماشین را صرفاً به‌عنوان دو فناوری مستقل یاد بگیرد. ارزش اصلی زمانی ایجاد می‌شود که این ابزارها در یک چارچوب مسئله‌محور شهرسازی قرار گیرند؛ یعنی ابتدا مسئله شهری تعریف شود، سپس داده‌های مناسب تهیه شوند، مدل انتخاب و ارزیابی شود و در نهایت خروجی الگوریتم به یک تحلیل فضایی و برنامه‌ریزی قابل تفسیر تبدیل شود.

به همین دلیل، یک پژوهش قوی در این حوزه باید به سه پرسش هم‌زمان پاسخ دهد:

۱. کجا احتمال توسعه شهری وجود دارد؟

۲. چه عوامل مکانی و غیرمکانی با این توسعه ارتباط دارند؟

۳. این پیش‌بینی چگونه می‌تواند به تحلیل و تصمیم‌سازی شهرسازی کمک کند؟

پاسخ هم‌زمان به این سه سؤال، تفاوت میان یک پروژه صرفاً فنی یادگیری ماشین و یک پژوهش کاربردی در مهندسی شهرسازی را مشخص می‌کند.

سوالات متداول

آیا GIS به تنهایی می‌تواند توسعه شهری آینده را پیش‌بینی کند؟

GIS عمدتاً یک چارچوب برای مدیریت، تحلیل و نمایش داده‌های مکانی است و به‌تنهایی الزاماً یک مدل پیش‌بینی‌کننده محسوب نمی‌شود. برای پیش‌بینی می‌توان GIS را با مدل‌های آماری، یادگیری ماشین، اتوماتای سلولی یا سایر روش‌های شبیه‌سازی ترکیب کرد.

برای پیش‌بینی توسعه شهری Random Forest بهتر است یا SVM؟

نمی‌توان بدون توجه به داده و طراحی پژوهش یکی را به‌طور کلی بهتر دانست. هر دو روش در مطالعات مکانی و طبقه‌بندی کاربرد دارند. مقایسه تجربی روی داده واقعی، همراه با معیارهای ارزیابی مناسب و اعتبارسنجی فضایی، روش علمی‌تری برای انتخاب مدل است.

آیا می‌توان از تصاویر ماهواره‌ای برای آموزش مدل استفاده کرد؟

بله. تصاویر ماهواره‌ای می‌توانند برای استخراج کاربری و پوشش زمین، تشخیص تغییرات و تولید برخی متغیرهای ورودی مدل استفاده شوند. ترکیب سنجش‌ازدور، GIS و یادگیری ماشین یکی از محورهای مهم پژوهش‌های جدید در تحلیل شهری است.

آیا مدل یادگیری ماشین می‌تواند آینده شهر را با قطعیت پیش‌بینی کند؟

خیر. مدل‌های یادگیری ماشین بر اساس داده‌های موجود الگو یاد می‌گیرند و خروجی آن‌ها به فرضیات، کیفیت داده و شرایط آینده وابسته است. بنابراین بهتر است نتایج به‌صورت احتمال، سناریو یا سطح پتانسیل توسعه تفسیر شوند، نه پیش‌بینی قطعی آینده.

برای پایان‌نامه شهرسازی چه ترکیبی از روش‌ها مناسب است؟

یک چارچوب قابل بررسی می‌تواند شامل GIS برای آماده‌سازی داده، سنجش‌ازدور برای استخراج تغییرات، Random Forest یا SVM برای مدل‌سازی و در صورت نیاز Cellular Automata برای شبیه‌سازی فضایی باشد. انتخاب نهایی باید بر اساس سؤال پژوهش و نوع داده انجام شود.

آیا یادگیری عمیق برای همه پروژه‌های توسعه شهری مناسب است؟

خیر. یادگیری عمیق زمانی توجیه بیشتری پیدا می‌کند که داده کافی، ساختار مناسب داده و مسئله‌ای با پیچیدگی لازم وجود داشته باشد. برای بسیاری از مجموعه‌داده‌های جدولی، مدل‌های کلاسیک مانند Random Forest یا مدل‌های آماری ممکن است انتخاب عملی‌تر و قابل‌تفسیرتری باشند.

مهم‌ترین خطای روش‌شناختی در این نوع پژوهش چیست؟

یکی از خطاهای مهم، اعتبارسنجی نادرست داده‌های مکانی است. اگر نمونه‌های آموزش و آزمون از نظر مکانی بیش از حد مشابه باشند، عملکرد مدل ممکن است غیرواقعی بالا برآورد شود. نشت اطلاعات و استفاده از متغیرهایی که در زمان پیش‌بینی در دسترس نبوده‌اند نیز از خطاهای جدی هستند.

منابع

۱. Dritsas, E., & Trigka, M. (۲۰۲۶). *Advances in geospatial artificial intelligence for remote sensing applications*. Computer Science Review, ۶۰, ۱۰۰۹۱۳.
۲. *A Review of Machine Learning Modeling Approaches of Spatiotemporal Urbanization and Land Use Land Cover*. ۲۰۲۶.
۳. *Machine Learning in Land Use Prediction: A Comprehensive Review of Performance, Challenges, and Planning Applications*. Buildings, ۲۰۲۵.
۴. *Machine Learning Algorithms for Urban Land Use Planning: A Review*. Urban Science, ۲۰۲۱.
۵. *Cellular Automata in Modeling and Predicting Urban Densification: Revisiting the Literature since 1971*. Land, ۲۰۲۲, ۱۱(۷), ۱۱۱۳.
۶. *Scenario simulation studies of urban development using remote sensing and GIS: review*. Remote Sensing Applications: Society and Environment, ۲۰۲۱, ۱۰۰۴۷۴.
۷. *Urban Growth Forecast Using Machine Learning Algorithms and GIS-Based Novel Techniques: A Case Study Focusing on Nasiriyah City, Southern Iraq*. ISPRS International Journal of Geo-Information, ۲۰۲۳.
۸. *Machine learning and remote sensing integration for leveraging urban sustainability: A review and framework*. Sustainable Cities and Society.
۹. Google Earth Engine. *Machine Learning in Earth Engine*.