



فصلنامه علوم محیطی، دوره نوزدهم، شماره ۴، زمستان ۱۴۰۰

۱۸۳-۲۰۴

مقاله پژوهشی

شبیه سازی رشد شهری با استفاده از مدل سلول های خودکار و الگوریتم های یادگیری ماشین، مطالعه موردی: کلانشهر تبریز

امید اشکریز، بابک میرباقری*، علی اکبر متکان و علیرضا شکیبا

مرکز مطالعات سنجش از دور و GIS، دانشکده علوم زمین، دانشگاه شهید بهشتی، تهران، ایران

تاریخ دریافت: ۱۳۹۹/۹/۳ تاریخ پذیرش: ۱۴۰۰/۶/۱۴

اشکریز، ا.، ب. میرباقری، ع.ا. متکان و ع.ر. شکیبا. ۱۴۰۰. شبیه سازی رشد شهری با استفاده از مدل سلول های خودکار و الگوریتم های یادگیری ماشین، مطالعه موردی: کلانشهر تبریز. فصلنامه علوم محیطی. ۱۹(۴): ۱۸۳-۲۰۴.

سابقه و هدف: روند رشد شهری در دهه های اخیر تسریع شده است. بنابراین پیش بینی الگوی رشد آینده شهر برای جلوگیری از برخی مشکل های محیط زیستی، اقتصادی و اجتماعی از اهمیت بالایی برخوردار است. شهر تبریز نیز به دلیل تغییرات جمعیتی قابل توجه از رشد سریع زمین های شهری برخوردار بوده و بنابراین نیازمند شبیه سازی دقیق رشد شهری برای جلوگیری از پیامدهای منفی محیط زیستی و اقتصادی می باشد. هدف این پژوهش، ارزیابی دقت الگوریتم های مطرح یادگیری ماشین با روش اعتبارسنجی متقابل مکانی و تلفیق آن ها با مدل سلول های خودکار جهت شبیه سازی رشد زمین های شهری است.

مواد و روش ها: در این پژوهش جهت تحلیل تغییرات کاربری زمین های شهری تصاویر ماهواره ای لندست مربوط به سال های ۱۳۷۶، ۱۳۸۵ و ۱۳۹۴ با استفاده از روش ماشین بردار پشتیبان طبقه بندی شدند. در گام بعدی تولید نقشه های پتانسیل تغییر زمین های غیر شهری به شهر با استفاده از الگوریتم های جنگل تصادفی، ماشین بردار پشتیبان و شبکه عصبی پرسپترون چند لایه برای دو دوره کالیبراسیون (سال های ۱۳۷۶ - ۱۳۸۵) و اعتبارسنجی (سال های ۱۳۸۵ - ۱۳۹۴) براساس لایه های اطلاعاتی فاصله از معیار اصلی، فاصله از مرکز شهر، فاصله از زمین های ساخته شده، فاصله از رودخانه و راه آهن و همچنین لایه های شیب، ارتفاع و لایه دو کلاسه کاربری شامل کاربری کشاورزی (با مقدار صفر) و بایر (با مقدار یک)، انجام شد. در پژوهش حاضر جهت جلوگیری از بیش برآزش الگوریتم ها به نمونه های آموزشی و در نتیجه به دست آمدن نتایج خوشبینانه، از روش اعتبارسنجی متقابل مکانی با هدف کاهش همبستگی مکانی میان داده های آموزشی و آزمایشی، در فرایند استخراج پارامترهای بهینه الگوریتم های یادگیری ماشین استفاده گردید. در نهایت با استفاده از مدل سلول های خودکار، شبیه سازی رشد شهر تبریز براساس نقشه های کاربری زمین ها و پتانسیل تغییر به دست آمده از الگوریتم های یادگیری ماشین برای دوره های بیان شده انجام شد.

نتایج و بحث: نتایج نشان داد الگوریتم جنگل تصادفی با مقدار مساحت زیر منحنی ROC معادل ۰/۹۲۲۸ نسبت به الگوریتم های ماشین بردار پشتیبان و شبکه عصبی مصنوعی به ترتیب با مساحت های ۰/۸۹۵۱ و ۰/۸۷۲۶، عملکرد بهتری در برآورد پتانسیل تغییر زمین های

* Corresponding Author: Email Address. b_mirbagheri@sbu.ac.ir

<http://dx.doi.org/10.52547/envs.2021.36900>

<http://dorl.net/dor/20.1001.1.17351324.1400.19.4.12.0>

غیر شهری به شهر داشته است. همچنین این الگوریتم در مقایسه با سایر الگوریتم‌ها تغییرات محلی در پتانسیل تغییر را به صورت بارزتری استخراج کرده است. در نهایت رشد شهر تبریز با استفاده از مدل سلول‌های خودکار بر مبنای نقشه‌های پتانسیل تغییر به دست آمده شبیه‌سازی شد. مقایسه نقشه پیش‌بینی در دوره اعتبارسنجی با وضع موجود زمین‌های شهری در سال ۱۳۹۴ نشان داد که شبیه‌سازی رشد شهری با مدل سلول‌های خودکار مبتنی بر جنگل تصادفی با مقدار سنجه سازگاری $0/3569$ نسبت به مدل‌های مبتنی بر بردار پشتیبان و شبکه عصبی به ترتیب با مقادیر سنجه سازگاری $0/3496$ و $0/3434$ در اختصاص زمین‌های غیرشهری به شهری دقیق‌تر بوده است.

نتیجه‌گیری: توانایی الگوریتم‌های یادگیری ماشین نظیر شبکه‌های عصبی مصنوعی، ماشین بردار پشتیبان و جنگل تصادفی در حل مسئله‌های غیر خطی، استفاده از آن‌ها را در شبیه‌سازی رشد زمین‌های شهری اجتناب ناپذیر می‌سازد. در این میان در تحقیق حاضر الگوریتم جنگل تصادفی که اساس آن بر یادگیری جمعی استوار است از مزیت بالاتری نسبت به دو الگوریتم ماشین بردار پشتیبان و شبکه عصبی مصنوعی برخوردار بوده است.

واژه‌های کلیدی: شبیه‌سازی رشد شهری، مدل سلول‌های خودکار، جنگل تصادفی، ماشین بردار پشتیبان، شبکه عصبی مصنوعی

مقدمه

به دست آمده به استراتژی‌های مدیریت شهری کمک می‌کند (Li and Gong, 2016). شبیه‌سازی رشد شهری در آینده با دقت بالا از جمله مهمترین چالش‌های مطالعات شهری به شمار می‌رود (Aburas et al., 2017) و بنابراین لازم است مدل‌های مناسبی برای پیش‌بینی رشد شهری مورد استفاده قرار گیرد.

از جمله مدل‌هایی که بیشتر در شبیه‌سازی فرایندهای پویای شهری مورد استفاده قرار می‌گیرد، مدل سلول‌های خودکار^۱ است (Yao et al., 2017). از دهه ۱۹۹۰، مدل سلول‌های خودکار توسط بسیاری از محققان در سراسر دنیا به منظور شبیه‌سازی رشد شهری به کار گرفته شده است، نظیر: Clarke et al. (1997); Wu and Webster, (1998); Batty and Xie, (2005); Mundia et al. (2007); Liu et al. (2014); Liang et al. (2018); Zhang et al. (2020). مدل سلول‌های خودکار یک رویکرد پایین به بالا در شبیه‌سازی سیستم‌های پیچیده، مانند شبیه‌سازی رشد شهری را دارا می‌باشد (Yao et al., 2017). استفاده از مدل سلول‌های خودکار در شبیه‌سازی رشد شهری به دلیل سادگی، انعطاف‌پذیری و بویژه توانایی بالا در ترکیب ابعاد مکانی و زمانی گسترش یافته است (He et al., 2018)؛ و با توجه به انطباق‌پذیری آن با GIS و سنجش از دور، علاقه بسیاری از محققان را

رشد شهرنشینی و صنعتی شدن سبب به وجود آمدن تغییرات سریع کاربری و پوشش زمین‌ها در دو قرن گذشته شده است (Jat et al., 2017). رشد شهرها موجب پیشرفت جوامع انسانی می‌شود و اقتصاد ملی را تقویت می‌کند. با این وجود، تعداد بسیاری از شهرها با سرعت بیش از حد رشد می‌کنند؛ که رشد بی‌رویه جمعیت و مهاجرت روستاییان به شهرها، سبب پیدایش و تداوم بحران‌های عظیم محیط زیستی و کاهش کیفیت زندگی جوامع شهری شده است (Maleki, 2010). براساس داده‌های سازمان ملل متحد، بیش از ۵۰٪ از جمعیت جهان در حال حاضر در منطقه‌های شهری زندگی می‌کنند و پیش‌بینی می‌شود تا سال ۲۰۵۰ این مقدار به ۷۲٪ افزایش یابد (UNFPA, 2016). اگر این روند ادامه یابد، زمین تبدیل شده به منطقه‌های شهری در ۲۰ سال آینده کمابیش سه برابر خواهد شد (Luederitz et al., 2013). از این رو مدل‌سازی رشد شهری بویژه در کلانشهرها، با هدف برنامه‌ریزی دقیق برای آینده از اهمیت بالایی برخوردار است (Hosseinali et al., 2013). مدل‌سازی رشد شهری نه تنها روند شهرنشینی را مورد بررسی قرار می‌دهد، بلکه منجر به شناخت نیروهای مؤثر بر رشد شهری و پیامدهای محیط زیستی آن نیز می‌شود و در نهایت با فراهم آوردن پشتیبانی قوی از اطلاعات

مدل نشان داد که به کارگیری الگوهای تلفیقی می‌تواند سبب افزایش دقت در فرایند کالیبراسیون قوانین انتقال مدل سلول‌های خودکار شود.

Yang *et al.* (2008) در پژوهش خود، از الگوریتم ماشین بردار پشتیبان برای تعریف قوانین انتقال مدل سلول‌های خودکار جهت پیش بینی رشد شهر شنژن^۶ چین استفاده کردند. آن‌ها توانایی مدل سلول‌های خودکار را در مواجهه با پیچیدگی‌های مسئله‌های غیرخطی از طریق انتقال داده‌های ورودی به یک فضای ویژگی^۷ با ایجاد ابر صفحه^۸ که در آن می‌توان طبقه‌بندی بهینه را از طریق به حداکثر رساندن حاشیه بین دو کلاس انجام داد، بهبود بخشیدند. نتایج نشان داد که این مدل دقت بالاتری نسبت به مدل رگرسیون لجستیک در شبیه‌سازی رشد شهری با سلول‌های خودکار دارد.

Mustafa *et al.* (2018) از رویکرد تلفیقی مدل سلول‌های خودکار و ماشین بردار پشتیبان برای شبیه‌سازی تغییر کاربری زمین‌های شهری در والونیا^۹ بلژیک بین سال‌های ۲۰۰۰ تا ۲۰۱۰ استفاده کردند. در این پژوهش برای ارزیابی دقت تخصیص زمین‌های جدید شهری، مدل ماشین بردار پشتیبان با مدل رگرسیون لجستیک مورد مقایسه قرار گرفت که نتایج نشان داد مدل ماشین بردار پشتیبان توانایی بهتری نسبت به مدل رگرسیون لجستیک در کالیبراسیون مدل سلول‌های خودکار دارد.

Zakir Haghghi and Haji Babaei (2013) در پژوهشی به شبیه‌سازی رشد شهری همدان از سال ۱۳۳۵ تا سال ۱۳۹۱ با استفاده از مدل سلول‌های خودکار و روش ماشین بردار پشتیبان پرداختند که نتایج حاصل نشان داد مطابقت خوبی بین الگوهای رشد واقعی و شبیه‌سازی شده می‌باشد. آن‌ها نتیجه مدل سلول‌های خودکار براساس قوانین انتقال ماشین بردار پشتیبان را به‌عنوان یک ابزار مفید برای شبیه‌سازی رشد شهری معرفی کردند.

همچنین در زمینه استفاده از الگوریتم جنگل تصادفی می‌توان به تحقیق Kamusoko and Gamba (2015)

در زمینه مدل‌سازی رشد شهری به خود جلب کرده است (Sakieh *et al.*, 2014). هسته اصلی مدل سلول‌های خودکار چگونگی تعریف قوانین انتقال^۲ است به گونه‌ای که بر مبنای آن بتوان با به کارگیری داده‌های تجربی، الگوهای واقع‌گرایانه‌ای را شبیه‌سازی نمود (Liu *et al.*, 2008).

رشد شهری فرایند پیچیده و غیر خطی است و مدل سلول‌های خودکار نیز یک مدل پویا و دارای مزیت‌های منحصر به فرد در شبیه‌سازی مسئله‌های پیچیده به‌شمار می‌رود (Yeh *et al.*, 2021). از طرف دیگر الگوریتم‌های یادگیری ماشین که از آن‌ها در جهت تشخیص روابط غیرخطی بین پدیده‌ها استفاده می‌شود، می‌توانند در حل مسئله‌های پیچیده غیر خطی شهری کارآمد باشند (Yang *et al.*, 2008). از میان الگوریتم‌های زیاد یادگیری ماشین، شبکه‌های عصبی مصنوعی^۳، ماشین بردار پشتیبان^۴ و جنگل تصادفی^۵ از جمله الگوریتم‌های مرسوم بکار گرفته شده در شبیه‌سازی تغییرات کاربری زمین‌ها در تلفیق با مدل سلول‌های خودکار به‌شمار می‌روند (Qian, 2020).

از میان تحقیق‌های انجام شده با استفاده از الگوریتم‌های بالا می‌توان به تحقیق Samadzadegan and Kiavarz (2009) اشاره نمود که در پژوهش خود، الگوی تلفیقی سلول‌های خودکار و شبکه عصبی مصنوعی را برای شبیه‌سازی تغییرات کاربری زمین‌ها در شهر اصفهان پیشنهاد کردند. هدف آن‌ها، آزمودن الگوی یاد شده برای شبیه‌سازی تغییرات کاربری زمین‌ها بود و به این منظور، چهار دهه تغییرات را از سال‌های ۱۹۷۰ تا ۲۰۰۱ بررسی کردند.

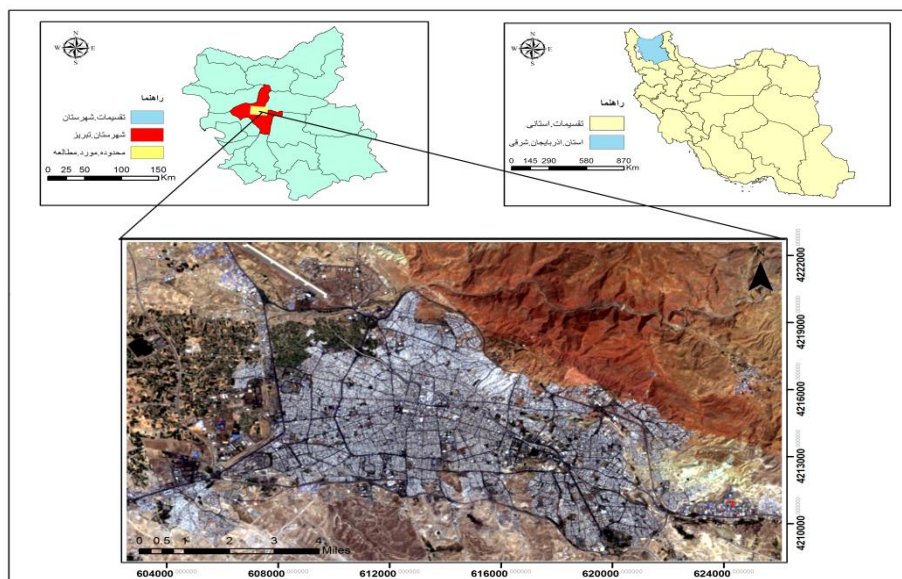
Salehi Asfichi *et al.* (2018) در پژوهشی به ارزیابی کارایی الگوی تلفیقی مدل سلول‌های خودکار و شبکه عصبی مصنوعی در شبیه‌سازی رشد شهری کلانشهر تهران از سال ۱۹۹۰ تا سال ۲۰۱۰ پرداختند. مقایسه آماری نقشه واقعیت زمینی سال ۲۰۱۰ با نقشه حاصل از

مواد و روش‌ها

منطقه مورد مطالعه و داده‌های تحقیق

کلانشهر تبریز، مرکز استان آذربایجان شرقی، بزرگترین شهر شمال غرب ایران است که با داشتن ۱ میلیون ۷۷۳ هزار و ۱۳۳ نفر بیش از ۴۰ درصد از جمعیت استان را در خود جای داده است و ششمین شهر پر جمعیت ایران شناخته می‌شود (Statistical Center of Iran, 2016). از لحاظ مختصات جغرافیایی، شهر تبریز در ۴۶ درجه و ۱۷ دقیقه شرقی و ۳۸ درجه و ۵ دقیقه عرض شمالی در ارتفاع تقریبی ۱۳۵۰ متر از سطح دریا واقع شده است که از سمت شمال شرق به کوه‌های عون ابن علی، از جنوب به رشته کوه‌های سه‌سند و از غرب با شیب ملایمی به دریاچه ارومیه ختم می‌شود. طی دهه‌های اخیر، شهر تبریز دستخوش یک رشد نامنظم و سریع شده است و تغییرات جمعیتی چشم‌گیری را تجربه کرده است که با توجه با موقعیت جغرافیایی این شهر، در برابر مشکل‌های ناشی از رشد غیرمترقبه شهری آسیب پذیرتر شده است (Moosavi, 2011). شکل ۱ موقعیت جغرافیایی شهر تبریز را نمایش می‌دهد.

اشاره نمود که در آن به شبیه‌سازی رشد شهر هراره^{۱۰} پایتخت کشور زیمبابوه با استفاده از مدل سلول‌های خودکار و الگوریتم جنگل تصادفی پرداختند. آن‌ها همچنین برای ارزیابی دقت شبیه‌سازی، این الگوریتم را با الگوریتم‌های ماشین بردار پشتیبان و رگرسیون لجستیک مورد مقایسه قرار دادند. نتایج نشان داد تلفیق الگوریتم جنگل تصادفی و مدل سلول‌های خودکار از پتانسیل بالایی در شبیه‌سازی رشد شهری برخوردار است. همانطور که مرور پیشینه تحقیقاتی نشان می‌دهد، الگوریتم‌های شبکه عصبی مصنوعی، ماشین‌های بردار پشتیبان و جنگل‌های تصادفی در موارد زیادی جهت شبیه‌سازی رشد زمین‌های شهری با استفاده از مدل سلول‌های خودکار بکار گرفته شده‌اند. هدف تحقیق حاضر مقایسه نتایج حاصل از بکارگیری این سه الگوریتم در پیش‌بینی رشد زمین‌های شهری و تعیین مدل بهینه است. همچنین در تحقیق حاضر برخلاف تحقیقاتی مشابه، به جهت کاهش اثرهای خودهمبستگی مکانی و پرهیز از به‌دست آمدن دقت‌های خوشبینانه، در تعیین پارامترهای بهینه هر یک از الگوریتم‌های بالا از روش اعتبارسنجی متقابل مکانی^{۱۱} استفاده شده است.



شکل ۱- موقعیت جغرافیایی منطقه مورد مطالعه
Fig. 1- Geographical location of the study area

پژوهش، تصویرهای لندست^{۱۴} و ۵ و ۸ با قدرت تفکیک مکانی ۳۰ متر می‌باشد که برای ارزیابی میزان گسترش زمین‌های شهری و استخراج طبقات کاربری زمین‌های سال-های ۱۳۷۶، ۱۳۸۵ و ۱۳۹۴ مورد استفاده قرار گرفته است. جدول ۱، ویژگی‌های مربوط به تصویرهای ماهواره‌ای لندست را نشان می‌دهد.

داده‌های پایه مورد استفاده در این پژوهش شامل تصاویر ماهواره‌ای، مدل رقومی ارتفاع^{۱۲} و داده‌های برداری مربوط به شبکه راه‌ها (بزرگراه‌ها، آزادراه‌ها و خیابان‌ها)، رودخانه و خط‌های راه‌آهن می‌باشد. این داده‌ها از سازمان نقشه برداری کشور و سایت سازمان زمین شناسی آمریکا (USGS)^{۱۳} به-دست آمد. تصویرهای ماهواره‌ای استفاده شده در این

جدول ۱- ویژگی‌های تصاویر ماهواره‌ای لندست

Table 1. Specifications of Landsat satellite images

شماره تصویر	ماهواره	سنجده	تاریخ اخذ تصویر	شماره ردیف و گذر	مکانی	قدرت تفکیک
	Satellite	Sensor	Date of acquisition	Row and path	Spatial resolution	Radiometric resolution
1	Landsat5	TM	1997-8-15	168-03	30 Meter	8 Bits
2	Landsat5	TM	2006-9-9	168-03	30 Meter	8 Bits
3	Landsat8	OLI	2015-7-16	168-03	30 Meter	16 Bits

آبی، پارک‌ها و فضای سبز، زمین‌های کشاورزی و باغ‌ها و منطقه‌های بایر استخراج شد. شکل ۲، نقشه‌های کاربری زمین‌های منطقه مورد مطالعه را نشان می‌دهد. برای ارزیابی دقت طبقه‌بندی تصویرهای ماهواره‌ای نمونه‌های واقعی گرفته شده از تصاویر Google Earth و نقشه‌های ۱:۲۰۰۰ شهر تبریز با نقشه‌های کاربری زمین‌های حاصل از طبقه‌بندی تصاویر ماهواره‌ای مقایسه شده و سپس برای هر یک از نقشه‌ها سنج‌های دقت کلی و کاپا محاسبه گردید. در جدول ۲، ویژگی‌های مربوط به نقشه‌های کاربری زمین‌ها و ارزیابی دقت آن‌ها نمایش داده شده است.

تهیه نقشه کاربری زمین‌ها و عامل‌های مؤثر در

رشد شهری

به‌منظور تهیه نقشه‌های کاربری زمین‌های شهر تبریز، تصاویر سنجنده‌های ماهواره لندست در ماه‌های بدون ابر و در بالاترین میزان سبزی‌نگی مورد استفاده قرار گرفت. از آنجا که تصاویرها در سال‌های متفاوتی برداشت شده است، تصحیح‌های رادیومتریک، اتمسفری و هندسی روی تصویرها اعمال گردید. سپس با استفاده از طبقه‌بندی نظارت شده و روش ماشین بردار پشتیبان نقشه‌های کاربری زمین‌های شهر تبریز با پنج کاربری عمده منطقه‌های شهری، پهنه‌های

جدول ۲- اطلاعات مربوط به نقشه‌های کاربری زمین‌ها و ارزیابی دقت آن‌ها

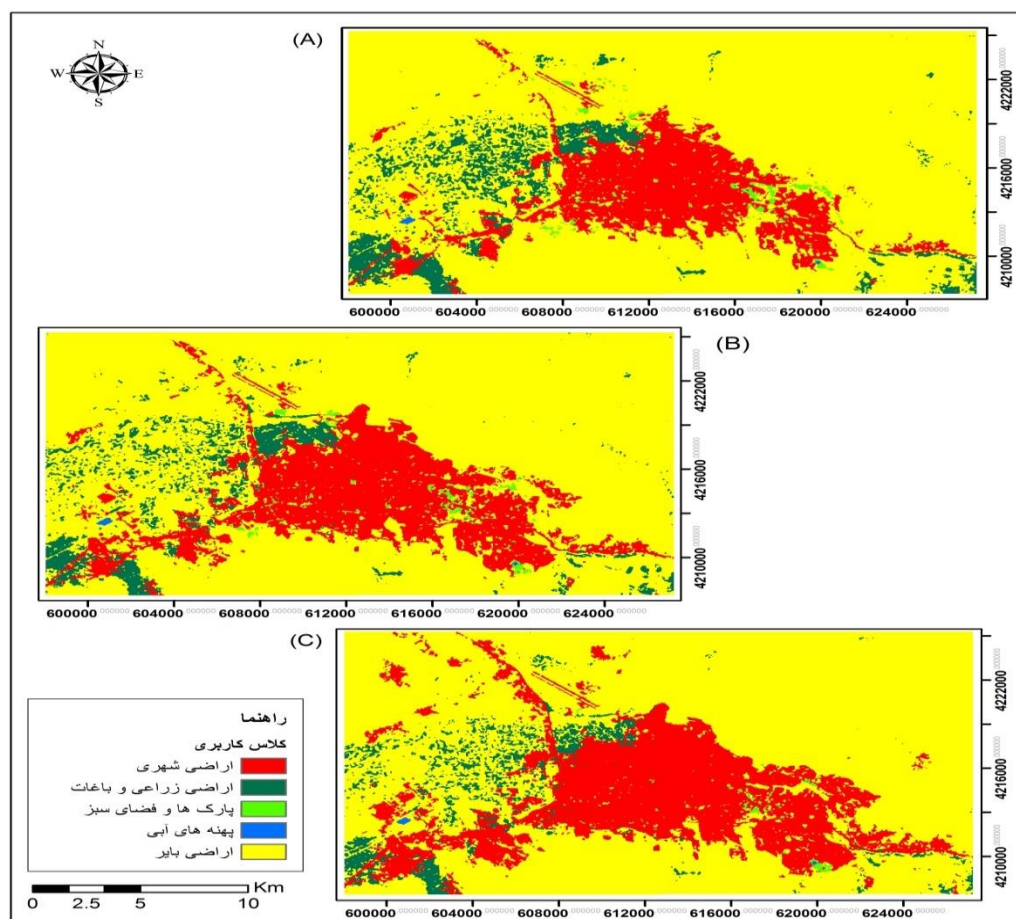
Table 2. Information on land use maps and assessing their accuracy

نقشه کاربری زمین‌ها ۱۳۹۴		نقشه کاربری زمین‌ها ۱۳۸۵		نقشه کاربری اراضی ۱۳۷۶		
Land use map 2015		Land use map 2006		Land use map 1997		
درصد	مساحت	درصد	مساحت	درصد	مساحت	
Percentage	Area (km ²)	Percentage	Area (km ²)	Percentage	Area (km ²)	
23.34	121.74	18.74	97.77	15.07	78.63	زمین‌های شهری Urban
0.32	1.67	0.68	3.55	0.79	4.12	فضای سبز و پارک‌ها Parks
0.04	0.23	0.05	0.26	0.05	0.27	پهنه آبی Water
4.49	24.45	5.68	29.64	6.95	36.24	زمین‌های زراعی و باغ‌ها Agricultural lands and gardens
71.59	373.36	74.83	390.24	77.12	402.21	زمین‌های بایر Barren lands

ادامه جدول ۲- اطلاعات مربوط به نقشه های کاربری زمین ها و ارزیابی دقت آن ها

Cont. Table 2. Information on land use maps and assessing their accuracy

نقشه کاربری زمین ها ۱۳۹۴ Land use map 2015	نقشه کاربری زمین ها ۱۳۸۵ Land use map 2006	نقشه کاربری اراضی ۱۳۷۶ Land use map 1997	
0.932	0.901	0.895	دقت کلی Overall accuracy
0.881	0.824	0.816	کاپا Kappa



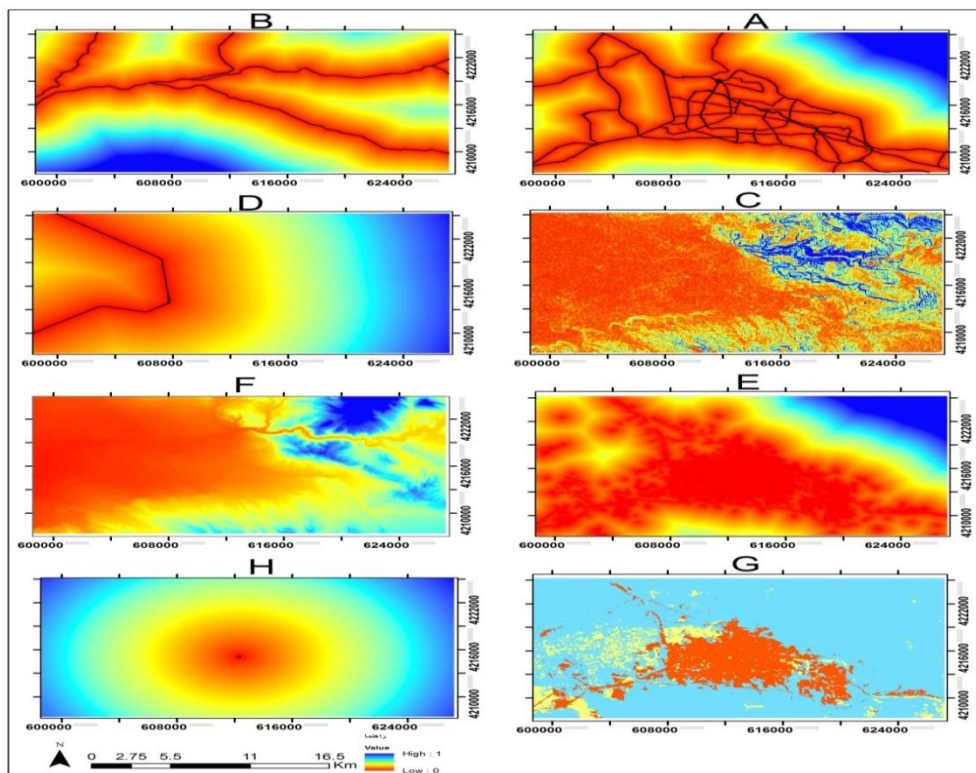
شکل ۲- نقشه کاربری زمین های منطقه مورد مطالعه به ترتیب برای سال های A: ۱۳۷۶، B: ۱۳۸۵ و C: ۱۳۹۴

Fig. 2- Land use map of the study area in A: 1997, B: 2006, and C: 2015

و ۰/۱۱ درصد از مساحت زمین های بایر، زمین های زراعی و باغ ها و فضای سبز و پارک ها به زمین های شهری تبدیل شده اند که زمین های شهری در طول دوره ۹ ساله با نرخ افزایش سالیانه ۲/۱۲ کیلومتر مربع، ۳/۶۷ درصد نسبت به سال ۱۳۷۶ رشد داشته است. در دوره دوم (اعتبارسنجی) نیز در مجموع ۵/۶ درصد از زمین های بایر، زمین های زراعی و باغ ها و فضای سبز و پارک ها به زمین های شهری تبدیل شده اند که زمین های شهری از سال

در مطالعه حاضر دوره زمانی (۱۳۷۶ - ۱۳۸۵) به عنوان دوره کالیبراسیون و دوره زمانی (۱۳۸۵ - ۱۳۹۴) به عنوان دوره اعتبارسنجی شبیه سازی رشد شهری انتخاب گردیدند. نرخ تغییرات زمین های شهری در محدوده مورد مطالعه برای دو دوره کالیبراسیون و اعتبارسنجی با استفاده از نقشه های کاربری زمین های سال های ۱۳۷۶، ۱۳۸۵ و ۱۳۹۴ به صورت مقایسه دو به دو محاسبه گردید. در دوره کالیبراسیون به ترتیب ۲/۲۹ درصد، ۱/۲۷ درصد

لایه دو کلاسه کاربری شامل کشاورزی (با مقدار صفر) و بایر (با مقدار یک)، به عنوان عامل‌های مؤثر در رشد شهر انتخاب شده و هریک از آن‌ها بین مقادیر صفر و یک استاندارد گردید. شکل ۳، نقشه‌های عامل‌های مؤثر در رشد شهر تبریز را در حالت استاندارد شده نشان می‌دهد.



شکل ۳- لایه‌های استاندارد شده مورد استفاده در شبیه‌سازی رشد شهری، A: فاصله از معابر اصلی، B: فاصله از رودخانه، C: ارتفاع، D: فاصله از خط‌های راه آهن، E: فاصله از زمین‌های ساخته شده، F: شیب، G: کاربری کشاورزی/بایر، H: فاصله از مرکز شهر
 Fig. 3- Standardized layers used in urban growth simulation. A: Distance from main roads. B: Distance from rivers. C: Height. D: Distance from railways. E: Distance from built-up areas. F: Slope. G: Land use (agriculture/barren). H: Distance from the city center.

در این بین، شبکه‌های عصبی پرسپترون چندلایه^{۱۵} که توسط (Rumelhart *et al.*, 1986) طراحی شدند از پرکاربردترین شبکه‌های عصبی مصنوعی می‌باشد که عمومی‌ترین ساختار این شبکه ساختار پیش‌رونده است که دارای سه لایه ورودی، میانی و خروجی می‌باشد و در آن ورودی هر نورون تنها از لایه قبل از آن به دست می‌آید (Sheikh Goodarzi *et al.*, 2016). وزن‌ها در شبکه عصبی مصنوعی به کمک الگوریتم‌های آموزشی برآورد و با به کارگیری توابع فعال‌سازی به نرون‌های خروجی منتقل می‌شوند (Pijanowski *et al.*, 2022). خروجی

۱۳۸۵ تا سال ۱۳۹۴ به میزان ۲۳/۹۷ کیلومتر مربع با نرخ افزایش سالیانه ۲/۶۶ کیلومتر مربع، رشد داشته است. در این پژوهش، لایه‌های فاصله از معابر اصلی، فاصله از مرکز شهر، فاصله از زمین‌های ساخته شده، فاصله از رودخانه و راه آهن و همچنین لایه‌های شیب، ارتفاع و

روش تحقیق

تولید نقشه‌های پتانسیل رشد زمین‌های شهری با استفاده از الگوریتم‌های یادگیری ماشین شبکه عصبی مصنوعی

از آنجاکه رشد شهری یک پدیده پیچیده است، تعداد زیادی از متغیرها به صورت غیر خطی با یکدیگر در تعامل هستند. در نتیجه به کارگیری شبکه‌های عصبی مصنوعی برای مدل‌سازی رشد شهری به طور کامل منطقی به نظر می‌رسد.

نهایی که متفاوت از مجموع سیگنال های دریافتی است، توسط توابع فعال سازی در بازه ای خاص محدود می شود. برای تولید نقشه پتانسیل رشد شهری از تابع فعال سازی Sigmoid که برد آن بین بازه صفر تا یک می باشد؛ استفاده می گردد (Javadi, 2008).

ماشین بردار پشتیبان

ماشین بردار پشتیبان از جمله روش های طبقه بندی نظارت شده است که برای تفکیک دو کلاس از هم با استفاده از داده های آموزشی مرز تصمیم گیری (ابرفصحه) بهینه را تعیین می کند. در این الگوریتم توازن میان حاشیه و خطای طبقه بندی توسط عدد ثابت C کنترل می شود. اگر طبقه بندی داده ها با استفاده از یک تابع خطی امکان پذیر نباشد، ماشین بردار پشتیبان می تواند با استفاده از توابع کرنل، به عنوان یک طبقه بندی کننده غیرخطی مورد استفاده قرار گیرد (Widodo *et al.*, 2007). اگرچه انواع مختلفی از توابع کرنل وجود دارد، تابع پایه شعاعی^{۱۶} (RBF) به عنوان بهترین انتخاب از بین دیگر توابع کرنل گزارش شده است (Yu *et al.*, 2006). از جمله پارامترهای این تابع پهنای باند (σ) می باشد که مقدار بهینه آن طی فرایند کالیبراسیون به همراه عدد ثابت C تعیین می گردد. به طور کلی ماشین بردار پشتیبان به عنوان یک طبقه بندی کننده دو کلاسه می تواند در استخراج پتانسیل تغییر مدل سلول های خودکار شهری مورد استفاده قرار گیرد (Ana *et al.*, 2004). پس از طبقه بندی می توان احتمال توسعه زمین های شهری را با استفاده از یک تابع Sigmoid (Platt, 1999; Wu *et al.*, 2004) محاسبه کرد.

جنگل تصادفی

الگوریتم جنگل تصادفی به عنوان یک روش غیر پارامتریک یادگیری ماشین متعلق به خانواده روش های یادگیری جمعی^{۱۷} است. این الگوریتم نوع توسعه یافته ای از الگوریتم طبقه بندی و رگرسیون درختی^{۱۸} است که توسط Brieman (2001) ارائه شده است. جنگل تصادفی

درختان تصمیم زیادی توسط روش نمونه برداری بوت استرپ^{۱۹} که بیشترین تمایز را با یکدیگر دارند تولید می کند (Ghasemi Esfahan, 2013). که برای انشعاب در هر گره، تنها از یکی از M ویژگی انتخاب شده به صورت تصادفی، استفاده می شود (Brieman, 2001). هر یک از درختان با توجه به ویژگی خود به یک کلاس رأی می دهد و در نهایت تخصیص هر نمونه آموزشی به یک کلاس بر مبنای بیشترین تعداد آرای که از طبقه بندی کننده ها دریافت می کنند، انجام می گیرد (Gislason *et al.*, 2006). در الگوریتم جنگل تصادفی نیاز به تعیین دو پارامتر مهم تعداد درخت های^{۲۰} تصمیم و تعداد زیر مجموعه تصادفی از ویژگی ها^{۲۱} می باشد (Huang and Boutros, 2016). پارامترهای مذکور روی خطای نهایی مؤثر است که باید مقادیر بهینه آن ها به دست آید (Pontius and Malanson, 2005). در الگوریتم جنگل تصادفی نیز احتمال توسعه زمین های شهری بر مبنای داده های ورودی قابل محاسبه است.

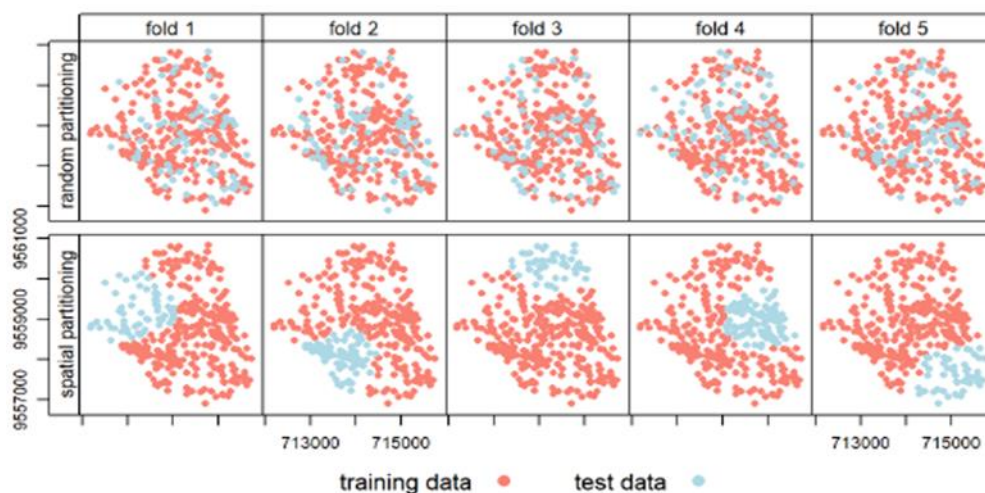
استخراج پارامترهای بهینه الگوریتم های

یادگیری ماشین به روش اعتبارسنجی متقابل

مکانی

تعیین مقدار بهینه پارامترهای مدل های یادگیری ماشین برای افزایش قدرت پیش بینی و مقایسه الگوریتم ها بسیار مهم است (Hutter *et al.*, 2011; Bergstra and Bengio, 2017; Duarte and Wainer, 2012). از طرف دیگر استفاده از روش اعتبارسنجی متقابل^{۲۲} معمولی با توجه به قانون اول جغرافیا (Tobler (1970) و شباهت بیشتر نقاط نزدیک به یکدیگر نسبت به نقاط دورتر سبب به وجود آمدن همبستگی مکانی بین نمونه های آموزشی و آزمایشی می گردد. این بدان معنی است که نمونه ها از نظر آماری مستقل نیستند، زیرا نمونه های آموزش و آزمایش در روش اعتبارسنجی متقابل معمولی به دلیل توزیع تصادفی نزدیک به یکدیگر هستند و ویژگی های مشابهی دارند که این مسئله منجر به بیش برآزش^{۲۳} مدل ها به

مکان آن‌هاست (Schratz *et al.*, 2018). همچنین این تقسیم‌بندی‌ها در دسته‌های مختلف تکرار می‌شوند تا تضمین کنند که هر نمونه یک بار در یکی از دسته‌های آزمایشی قرار گرفته باشد (Lovell *et al.*, 2019). شکل ۴ نحوه دسته بندی داده‌های آموزشی و آزمایشی را در روش-های اعتبارسنجی متقابل معمولی و مکانی نمایش می‌دهد.



شکل ۴- نحوه دسته‌بندی داده‌های آموزشی و آزمایشی در روش‌های اعتبارسنجی متقابل معمولی و مکانی

Fig. 4- Classifying training and test data in conventional and spatial cross-validation methods (Lovell *et al.*, 2019)

رشد شهری (سال‌های ۱۳۷۶ و ۱۳۸۵) براساس الگوریتم-های آموزش داده شده تولید گردید.

شبیه‌سازی رشد شهری با مدل سلول‌های خودکار

نوع پیکربندی مؤلفه‌های مدل سلول‌های خودکار، بر قوانین انتقال مدل تأثیر می‌گذارد (Lui, 2003). به‌طور کلی احتمال تبدیل هر سلول در مدل سلول‌های خودکار بر مبنای چهار مؤلفه اصلی محاسبه می‌گردد، که عبارتند از: پتانسیل تغییر (Pg)، همسایگی (Ω)، ضریب محدودیت (C) و فاکتور تصادفی (γ) (Feng *et al.*, 2016). پتانسیل تغییر (Pg): پتانسیل تغییر در مدل سلول‌های خودکار بیشتر براساس عامل‌های جغرافیایی و اقتصادی استوار است، و نحوه تعیین وزن و تأثیر عامل‌های مؤثر در رشد شهری چنان پیچیده است که به‌طور معمول نمی‌توان آن را با مدل‌های ساده ریاضی بیان کرد (He *et al.*, 2018). بنابراین در این پژوهش، پتانسیل تغییر زمین‌های

داده‌های آموزشی و به‌دست آمدن نتایج خوشبینانه می‌گردد (Breninng, 2012; Schratz *et al.*, 2018). بنابراین در این پژوهش، از روش اعتبارسنجی متقابل مکانی برای کاهش همبستگی مکانی بین نمونه‌های آموزشی و آزمایشی استفاده شده است و تفاوت آن با اعتبارسنجی متقابل معمولی، جداسازی داده‌های آموزشی و آزمایشی از یکدیگر براساس

در این پژوهش برای آموزش الگوریتم‌های شبکه عصبی پرسپترون چندلایه، ماشین بردار پشتیبان و جنگل تصادفی جهت استخراج پتانسیل رشد شهری از نرم افزار آماری متن باز R و بسته mlr (Bischl *et al.*, 2016) استفاده شد. ده درصد از مجموع نمونه‌های مربوط به زمین‌های رشد یافته و به همان تعداد زمین‌های رشد نیافته در دوره کالیبراسیون به‌عنوان نمونه آموزشی به-صورت تصادفی انتخاب گردید و از روش اعتبارسنجی متقابل مکانی به تعداد ۵ دسته و جستجوی تصادفی 2^4 به تعداد ۵۰۰ تکرار برای یافتن پارامترهای بهینه هر یک از الگوریتم‌ها استفاده شد. جستجوی تصادفی دارای ویژگی مناسب در ابعاد بالا است و در مقایسه با جستجوی شبکه-ای^{۲۵}، از دقت مناسبی برخوردار است (Bergstra and Bengio, 2012). در نهایت نقشه‌های پتانسیل تغییر زمین‌های شهری برای دوره‌های کالیبراسیون و اعتبارسنجی، با استفاده از لایه‌های عامل‌های مؤثر در

در شبیه سازی رشد شهری برای دوره های کالیبراسیون و اعتبارسنجی از سه مجموعه داده، شامل نقشه های کاربری زمین های سال های ۱۳۷۶ و ۱۳۸۵ به عنوان سال های پایه شبیه سازی، نقشه های پتانسیل تغییر در دوره های کالیبراسیون و اعتبارسنجی به دست آمده از الگوریتم های یادگیری ماشین و نرخ تغییرات رشد شهری دوره های کالیبراسیون و اعتبارسنجی استفاده شد. با توجه به بازه زمانی ۹ ساله دوره های کالیبراسیون و اعتبارسنجی، ۹ تکرار برای مدل سلول های خودکار در نظر گرفته شد. در هر تکرار، مدل سلول های خودکار تعداد مشخصی از سلول های غیر شهری بدون محدودیت را با بالاترین ضریب همسایگی شهری و پتانسیل تغییر (با در نظر گرفتن ضریب تصادفی)، براساس نرخ تغییرات دوره کالیبراسیون (۲۷۵۵ سلول) و اعتبارسنجی (۳۲۲۸ سلول) به سلول شهری تبدیل می کند.

ارزیابی دقت

منحنی مشخصه عملکرد نسبی (ROC)^{۲۷}

به منظور ارزیابی دقت نقشه های پتانسیل تغییر به دست آمده از الگوریتم های شبکه عصبی مصنوعی، ماشین بردار پشتیبان و جنگل تصادفی از منحنی ROC استفاده شد. محور عمودی این منحنی، نشانگر نرخ مثبت صحیح^{۲۸} (TPR) و محور افقی آن نشانگر نرخ مثبت کاذب (FPR)^{۲۹} است که به ترتیب از طریق رابطه های ۳ و ۴ محاسبه می شود (Kornac, 2006):

$$TPR = \frac{TP}{TP + FN} \quad (3)$$

$$FPR = \frac{FP}{FP + TN} \quad (4)$$

در روابط بالا، TP تعداد سلول های توسعه یافته شهری هستند که مدل آن ها را به درستی پیش بینی کرده، FN برابر سلول های توسعه یافته شهری است که مدل، قادر به پیش بینی آن ها نبوده است، FP برابر تعداد سلول های توسعه نیافته شهری هستند که مدل به اشتباه آن ها را به عنوان سلول توسعه یافته پیش بینی کرده و TN نیز برابر تعداد سلول های توسعه نیافته شهری هستند که مدل آن ها

شهری از الگوریتم های شبکه عصبی مصنوعی پرسپترون چندلایه، ماشین بردار پشتیبان و جنگل تصادفی که توانایی حل مسئله های غیر خطی را دارا می باشند، استخراج گردید. لازم به بیان است منظور از پتانسیل تغییر در این تحقیق، پتانسیل تغییر هر سلول از حالت غیر شهری به شهری یا همان پتانسیل رشد زمین های شهری است.

همسایگی (Ω): مجموعه ای از سلول هایی که با سلول مورد نظر اثر متقابل دارند، مورد بررسی قرار می گیرند، که یک عامل مهم در مدل سلول های خودکار است (He et al., 2018). همسایگی مور^{۲۶} با فیلتر ۳ در ۳ که به عنوان پرکاربردترین همسایگی در مدل سلول های خودکار شناخته می شود، در این پژوهش مورد استفاده قرار گرفت. ضریب محدودیت (C): نوع خاصی از کاربری زمین ها است که در طی فرایند شبیه سازی مجاز به تبدیل شدن به زمین های شهری نیست (Li and Yeh 2002; Wu and Martin, 2002) مانند: زمین های ساخته شده، پهنه های آبی، پارک ها و فرودگاه که در این تحقیق نیز به عنوان محدودیت در نظر گرفته شده اند. همچنین برای در نظر گرفتن عدم قطعیت در مدل سلول های خودکار، مؤلفه (γ) به عنوان یک ضریب تصادفی در مدل سلول های خودکار وارد می شود، که در این پژوهش به صورت رابطه ۱ تعریف شده است (White and Engelen, 1993).

$$\gamma = 1 + (-\ln(\text{rand}))^\alpha \quad (1)$$

که در آن rand متغیری تصادفی با توزیع یکنواخت در دامنه صفر و یک است و α پراکندگی توسعه شهری را کنترل می کند. در تحقیق حاضر شبیه سازی با مقادیر مختلف α از ۰.۱ تا ۱ به تعداد ۳۰ بار انجام شد و مقدار بهینه α تعیین گردید.

در نهایت احتمال تبدیل سلول i در زمان t به شهر در مدل سلول های خودکار براساس رابطه ۲ محاسبه می شود:

$$P_i^t = P g_i \times \Omega_i^t \times C_i \times \gamma_i \quad (2)$$

که در آن B، برابر تعداد سلول‌هایی است که کاربری آن‌ها در واقعیت تغییر کرده و مدل به درستی آن را پیش‌بینی کرده است. A برابر تعداد سلول‌هایی است که مدل کاربری آن‌ها را تغییر نداده، اما در واقعیت تغییر کاربری داده‌اند. C برابر تعداد سلول‌هایی است که توسط مدل به کاربری دیگری به غیر از شهری تغییر یافته است و D برابر تعداد سلول‌هایی است که در واقعیت تغییر نکرده‌اند اما توسط مدل تغییر یافته‌اند. از آنجاییکه در این پژوهش تغییرات کاربری زمین‌ها بر مبنای دو کلاس (شهری و غیر شهری) شبیه‌سازی می‌شود، مقدار پارامتر C برابر صفر خواهد بود (Yao et al., 2017).

نتایج و بحث

مقادیر بهینه پارامترهای الگوریتم‌های یادگیری

ماشین

جدول ۳ مقادیر بهینه پارامترهای هر یک از الگوریتم‌های یادگیری ماشین را که از روش جستجوی تصادفی با اعتبارسنجی متقابل مکانی به‌دست آمده است نمایش می‌دهد.

جدول ۳- مقادیر بهینه پارامترها

Table 3. Optimal values of parameters

مقدار بهینه Optimal valus	دامنه مقادیر Range of valus	پارامترها Parameters	الگوریتم‌ها Algorithms
3	0 :10	mtry	جنگل تصادفی
703	0:5000	Number of trees	Random forest
367	2^{-12} ; 2^{15}	C	ماشین بردار پشتیبان
0.000238	2^{-15} ; 2^6	σ	Support vector machine
34	0 :1000	Number of neurons	شبکه عصبی مصنوعی Artificial neural network

حاضر الگوریتم جنگل تصادفی تعمیم پذیری بالاتری نسبت به سایر الگوریتم‌ها در روش اعتبارسنجی متقابل مکانی دارد. شکل ۵ منحنی ROC مربوط به سه الگوریتم بالا را نمایش می‌دهد. استفاده از روش یادگیری جمعی در الگوریتم جنگل تصادفی که در نهایت منجر به کاهش واریانس برآورد می‌گردد را می‌توان به‌عنوان دلیلی بر بالاتر بودن دقت این الگوریتم برشمرد (James et al., 2013).

را به درستی پیش‌بینی کرده است. مساحت سطح زیر این منحنی AUC^{۳۰} نامیده می‌شود. در واقع مقدار AUC بیانگر احتمال این است که یک نمونه تصادفی انتخاب شده از مجموعه نمونه‌های موجود دارای پیش‌بینی صحیحی باشد. از این رو مقادیر بالای AUC نشان دهنده قابلیت اطمینان بالا از نقشه‌های پتانسیل تغییر تولید شده است (Mirbagheri and Alimohammadi, 2018). مقادیر AUC بین ۰/۵ تا ۱ متغیر هستند که به‌طور کلی مقادیر AUC بین ۰/۷ تا ۰/۸ بیانگر یک پیش‌بینی خوب، بین ۰/۸ تا ۰/۹ پیش‌بینی بسیار خوب و مقادیر بیشتر از ۰/۹ نشان دهنده بیش‌بینی عالی مدل است (Pradhan et al., 2014).

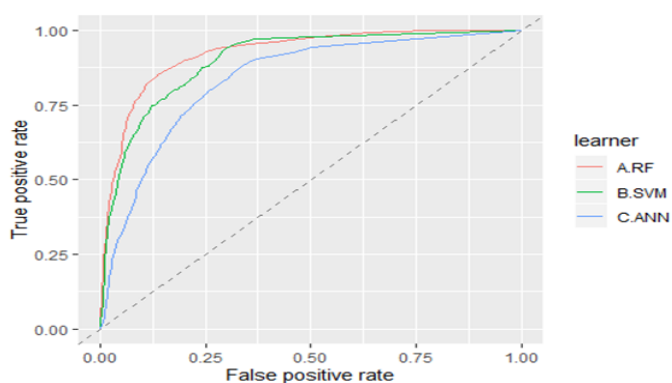
سنجه سازگاری

برای ارزیابی نتایج شبیه‌سازی رشد شهری از سنجه سازگاری^{۳۱} (FoM) که توسط Pontius et al. (2007) برای اولین بار پیشنهاد شد، استفاده گردید. دامنه مقادیر سنجه سازگاری بین صفر و یک متغیر است؛ که از طریق رابطه ۵ محاسبه می‌گردد.

$$FoM = \frac{B}{A + B + C + D} \quad (5)$$

ارزیابی دقت نقشه‌های پتانسیل رشد شهری

مقدار AUC برای الگوریتم‌های جنگل تصادفی، ماشین بردار پشتیبان و شبکه عصبی با پارامترهای بهینه به ترتیب برابر ۰/۸۷۲۶، ۰/۸۹۵۱، ۰/۹۲۲۸ و به‌دست آمد که نشان دهنده دقت عالی برای الگوریتم جنگل تصادفی و بسیار خوب برای الگوریتم‌های ماشین بردار پشتیبان و شبکه عصبی است. نتایج به‌دست آمده نشان می‌دهد در تحقیق



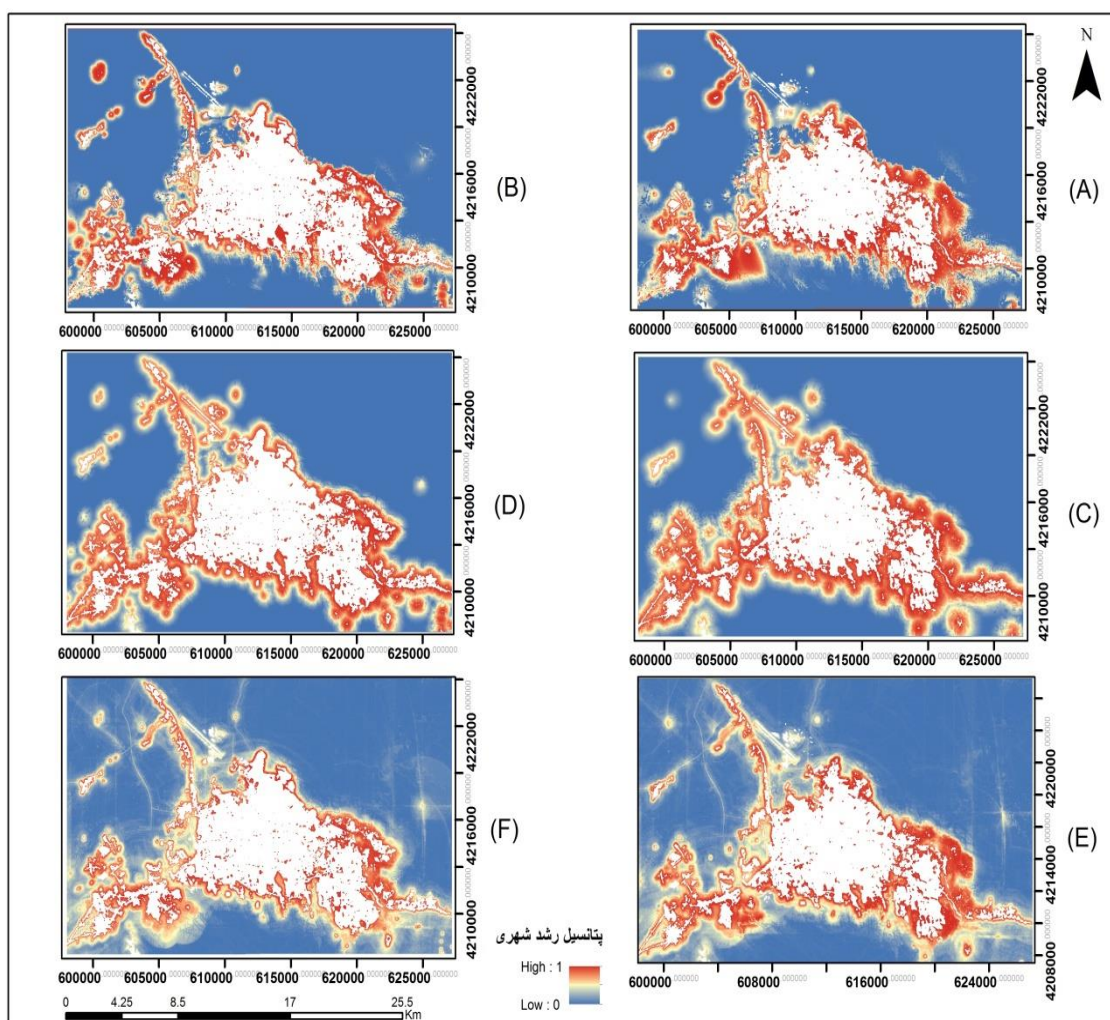
شکل ۵- نمودار مشخصه منحنی عملکرد نسبی

Fig. 5- Receiver operating characteristics diagram

الگوریتم های یادگیری ماشین برای دوره های کالیبراسیون و اعتبارسنجی را نشان می دهد.

نقشه های پتانسیل تغییر

شکل ۶ نقشه های پتانسیل تغییر حاصل از اجرای



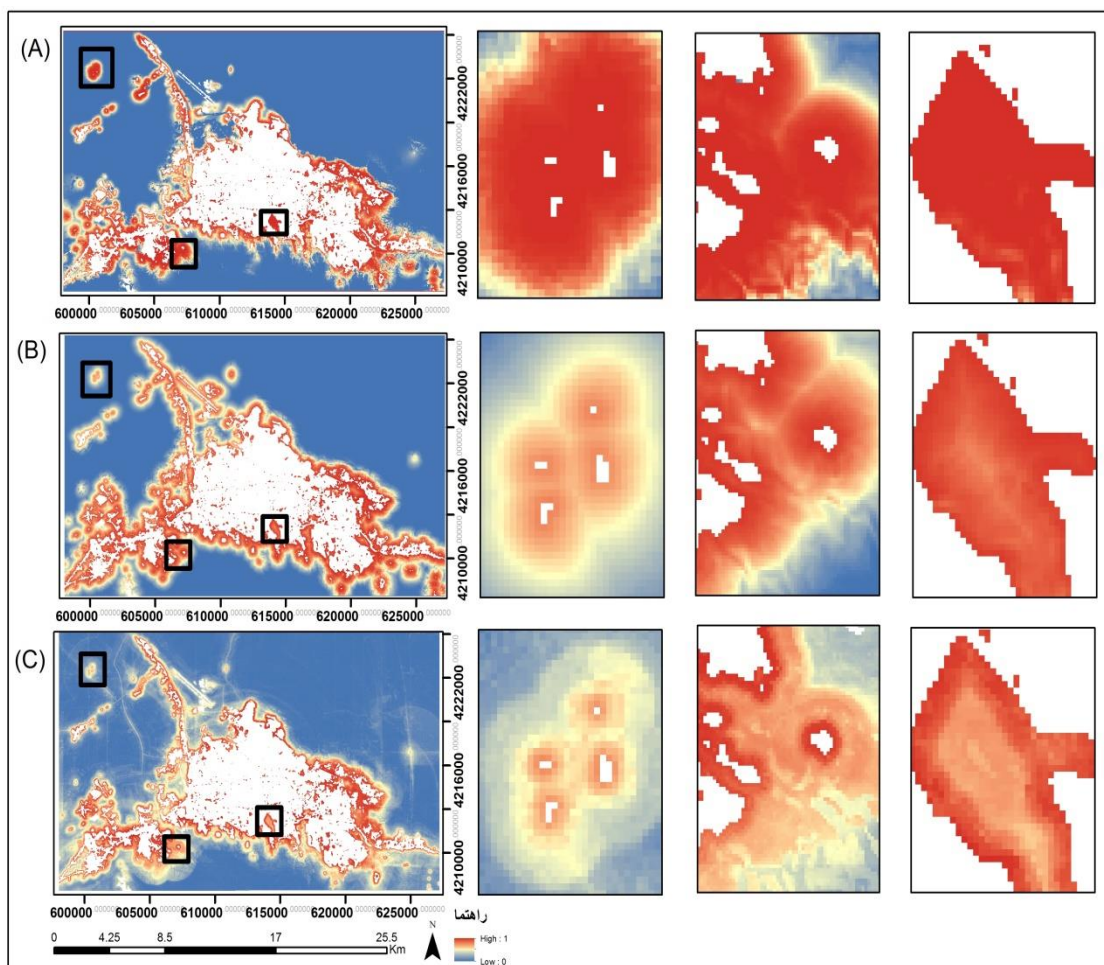
شکل ۶- نقشه های پتانسیل تغییر به ترتیب برای دوره های کالیبراسیون و اعتبارسنجی:

A و B: شبکه عصبی مصنوعی، C و D: ماشین بردار پشتیبان، E و F: جنگل تصادفی

Fig. 6- Change potential maps for calibration and validation periods, respectively. A and B: Artificial neural network. C and D: Support vector machine. E and F: Random forest

تغییر در این نقشه برابر ۰.۳۱ و در نقشه‌های حاصل از الگوریتم‌های ماشین بردار پشتیبان و شبکه عصبی مصنوعی به ترتیب برابر ۰/۲۸ و ۰/۲۶ است. شکل ۷ نقشه‌های پتانسیل تغییر دوره اعتبارسنجی را در بخش‌هایی از شهر با بزرگنمایی بیشتری نمایش می‌دهد. همانطور که در این شکل به خوبی مشاهده می‌شود، نقشه پتانسیل حاصل از پیاده‌سازی الگوریتم جنگل تصادفی دارای تغییرات محلی بیشتری نسبت به دو الگوریتم دیگر است.

تجزیه و تحلیل بصری نقشه‌های پتانسیل تغییر نشان می‌دهد که به‌طور کلی، الگوریتم‌های ماشین بردار پشتیبان و شبکه عصبی پرسپترون چندلایه پتانسیل تغییرات منطقه‌های غیر شهری به منطقه‌های شهری را به‌صورت یکنواخت‌تر برآورد کرده و تغییرات محلی^{۳۲} کمتری در آن‌ها مشاهده می‌شود. در مقابل در نقشه پتانسیل تغییر حاصل از اجرای الگوریتم جنگل تصادفی تغییرات محلی به‌صورت بارزتری مشاهده می‌شود به‌طوری‌که انحراف استاندارد مقادیر پتانسیل



شکل ۷- نقشه‌های پتانسیل تغییر بزرگ‌نمایی شده برای دوره اعتبارسنجی، A: شبکه عصبی مصنوعی،

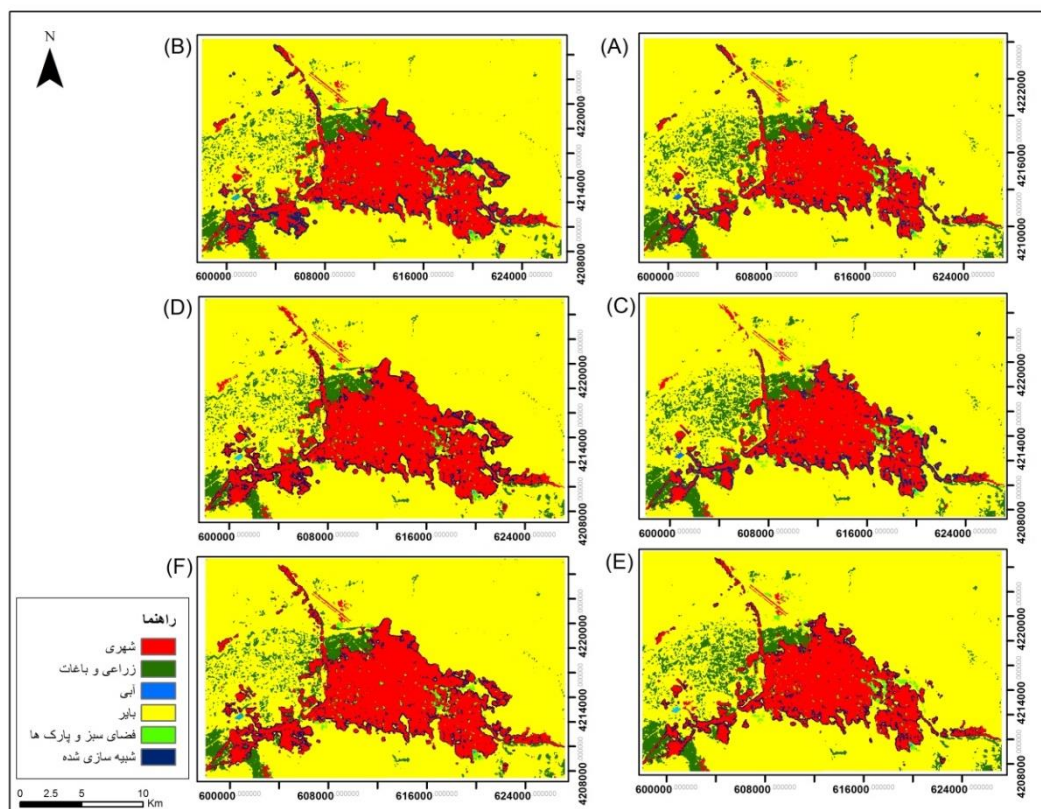
B: ماشین بردار پشتیبان، C: جنگل تصادفی

Fig. 7- Magnified change potential maps for the validation period. A: Artificial neural network. B: Support vector machine. C: Random forest

تغییر به‌دست آمده از الگوریتم‌های یادگیری ماشین، برای منطقه مورد مطالعه نشان می‌دهد. لازم به بیان است با توجه به شبیه‌سازی‌های انجام شده در مرحله کالیبراسیون، مقدار بهینه α برای ضریب تصادفی برابر ۰/۱ به‌دست آمد.

شبیه‌سازی رشد زمین‌های شهری

شکل ۸، نقشه‌های شبیه‌سازی رشد زمین‌های شهری حاصل از اجرای مدل سلول‌های خودکار را برای دوره‌های کالیبراسیون و اعتبارسنجی، با استفاده از نقشه‌های پتانسیل



شکل ۸ - نقشه های شبیه سازی رشد شهری به ترتیب برای دوره های کالیبراسیون و اعتبار سنجی، A و B: شبکه عصبی مصنوعی، C و D: ماشین بردار پشتیبان، E و F: جنگل تصادفی

Fig. 8- Urban growth simulation maps for calibration and validation periods, respectively. A and B: Artificial neural network. C and D: Support vector machine. E and F: Random forest

همانطور که پیشتر اشاره گردید، برای ارزیابی دقت شبیه سازی مدل سلول های خودکار در این مطالعه، از سنجه FoM استفاده شد که نتایج آن در جدول ۴ نمایش داده شده است.

جدول ۴- مقادیر FoM شبیه سازی ها در دو دوره کالیبراسیون و اعتبارسنجی
Table 4. FoM values of simulations in the two periods of calibration and validation

دوره اعتبارسنجی (۱۳۸۵-۱۳۹۴) Validation period (2006-2015)	دوره کالیبراسیون (۱۳۷۶-۱۳۸۵) Calibration period (1997-2006)	الگوریتم ها Algorithms
0.3569	0.3713	جنگل تصادفی Random forest
0.3496	0.3574	ماشین بردار پشتیبان Support vector machine
0.3434	0.3503	شبکه عصبی مصنوعی Artificial neural network

براساس مقایسه نقشه های شبیه سازی شده در دوره اعتبارسنجی با نقشه کاربری زمین های شهری سال ۱۳۹۴، مدل جنگل تصادفی توانسته است ۱۳۷۶۱۰۰۰ مترمربع از تغییرات کاربری زمین های شهری بین سال های ۱۳۸۵ تا ۱۳۹۴ را به طور صحیح پیش بینی کند؛ و

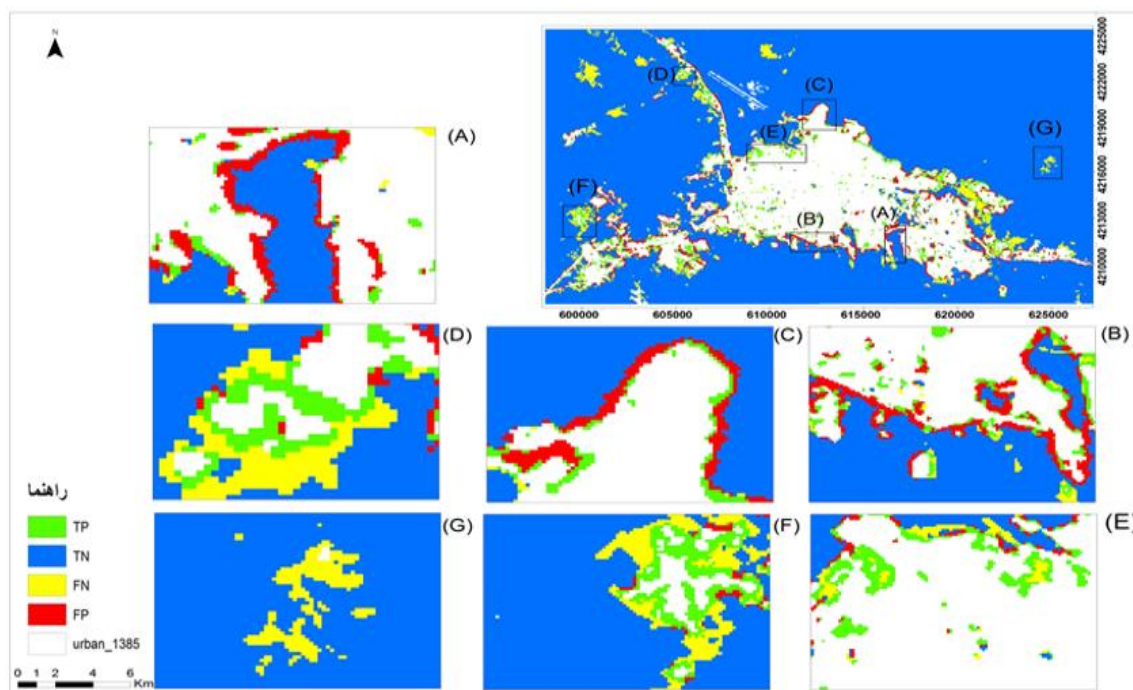
با توجه به نتایج به دست آمده از سنجه سازگاری الگوریتم جنگل تصادفی در مقایسه با الگوریتم های ماشین بردار پشتیبان و شبکه عصبی مصنوعی در شبیه سازی رشد شهری با استفاده از مدل سلول های خودکار در دوره های کالیبراسیون و اعتبارسنجی دقیق تر عمل کرده است.

لایه شبیه‌سازی شده با لایه واقعیت موجود مقایسه و براساس آن ماتریس در هم ریختگی^{۳۳} تشکیل می‌گردد که روشی مناسب برای مقایسه نقشه‌های مختلف از یک مکان در مقاطع زمانی متفاوت است (Asghari, 2015). شکل ۹ نقشه خطای محلی شبیه‌سازی رشد شهری با مدل سلول‌های خودکار مبتنی بر الگوریتم جنگل تصادفی را در سال ۱۳۹۴ نشان می‌دهد. در این شکل TP, FP, TN و FN مطابق با تعریف‌های ارائه شده در بخش ۱.۴.۳ هستند.

همچنین الگوریتم‌های ماشین بردار پشتیبان و شبکه عصبی مصنوعی با مساحت پیش‌بینی شده ۱۳۵۴۵۹۰۰ و ۱۳۳۷۴۹۰۰ مترمربع به ترتیب بعد از الگوریتم جنگل تصادفی از لحاظ دقت شبیه‌سازی رشد شهری قرار می‌گیرند.

تحلیل محلی نتایج

برای تحلیل نتایج و ارزیابی محلی نقشه‌های شبیه‌سازی شده رشد شهری، از مقایسه سلول مینا استفاده شد. در این روش



شکل ۹- نقشه خطای محلی شبیه‌سازی رشد شهری
Fig. 9- Local error map of urban growth simulation

منطقه‌ها و به‌طور کلی در حاشیه زمین‌های ساخته شده رخ داده است. در مقابل، رشد درونی شهر دقیق‌تر شبیه‌سازی شده است و کمترین میزان خطای FP در منطقه‌های دارای رشد درونی است (شکل ۹: D, E, F). همچنین برخی منطقه‌هایی مانند شهرک‌های مسکونی (شکل ۹: G) که رشد آن‌ها لبه‌ای و درونی نبوده، در واقع جزء زمین‌هایی هستند که به یک باره به منطقه‌های شهری تبدیل شده‌اند. در این منطقه‌ها مدل سلول‌های خودکار به دلیل ارزش پایین پتانسیل تغییر و کاهش ضریب همسایگی به خوبی نتوانسته است، رشد زمین‌های شهری را شبیه‌سازی کند.

به‌طور کلی گسترش زمین‌های شهری در سه قالب رشد بیرونی، رشد لبه‌ای و رشد درونی صورت می‌گیرد (Liu et al., 2010). در این بازه زمانی، بخش‌های شمالی و جنوبی شهر از سال ۱۳۸۵ تا ۱۳۹۴ به دلیل همجواری با شیب تند کوه‌ها (شکل ۹: C و B) و زمین‌های بایر تحت تملک دانشگاه تبریز (شکل ۹: A) رشد لبه‌ای قابل توجهی نداشته‌اند. از آنجایی که الگوریتم جنگل تصادفی بیشتر متأثر از عامل‌های فاصله از زمین‌های ساخته شده و شبکه معابر بوده است، در نتیجه این منطقه‌ها دارای ارزش بالایی در نقشه پتانسیل تغییر بوده و بیشترین خطای FP در این

مقایسه با سایر تحقیق‌ها

به‌طور کلی در تحقیق‌های به‌نسبت محدودی، کارایی انواع الگوریتم‌های یادگیری ماشین در شبیه‌سازی رشد شهری با استفاده از مدل CA با یکدیگر مقایسه شده است. با این حال در تحقیق‌های Shafizadeh-Moghadam *et al.* (2017) و Qian *et al.* (2020) دقت مدل CA مبتنی بر شبکه‌های عصبی مصنوعی از دقت مدل‌های CA مبتنی بر الگوریتم‌های جنگل تصادفی و ماشین بردار پشتیبان بالاتر بوده است. البته در تحقیق دیگری که توسط (He *et al.* (2018) به انجام رسیده است؛ دقت مدل CA مبتنی بر شبکه عصبی از دقت مدل CA مبتنی بر جنگل تصادفی در مرحله کالیبراسیون پایین‌تر و در مرحله اعتبار سنجی بالاتر بوده است. از طرف دیگر Kamusoko and Gamba (2015)، Qian, Shafizadeh-Moghadam *et al.* (2017) و Qian *et al.* (2020) نیز در تحقیق‌های خود مشابه تحقیق حاضر به این نتیجه رسیدند که دقت شبیه‌سازی رشد شهری مبتنی بر الگوریتم جنگل تصادفی و مدل سلول‌های خودکار دقت بالاتری نسبت به مدل CA مبتنی بر الگوریتم ماشین بردار پشتیبان دارد.

دلیل اختلاف بوجود آمده در نتایج مربوط به مدل CA مبتنی بر شبکه عصبی مصنوعی در این تحقیق و سایر تحقیقات مشابه می‌تواند ناشی از عامل‌های مختلفی نظیر تعداد لایه‌های پنهان مورد استفاده در شبکه عصبی (در تحقیق حاضر شبکه عصبی با یک لایه پنهان اجرا گردید)، نحوه کالیبراسیون مدل (در این تحقیق بر خلاف بسیاری از تحقیق‌های مشابه از روش اعتبارسنجی متقابل مکانی جهت کالیبراسیون مدل‌های یادگیری ماشین استفاده شد) و همچنین تفاوت در منطقه مطالعاتی و متغیرهای مورد استفاده در مدل‌سازی باشد. وجود تفاوت در منطقه‌های مطالعاتی که بستر اجرای الگوریتم‌های یادگیری ماشین هستند؛ خود به تنهایی می‌تواند عامل مهمی باشد در اینکه یک الگوریتم خاص همواره بهترین جواب را در مقایسه با سایر الگوریتم‌ها نداشته باشد.

نتیجه‌گیری

شبیه‌سازی رشد شهری می‌تواند اطلاعات مهمی از جمله میزان، موقعیت و زمان تغییرات زمین‌های شهری را برای دست‌اندرکاران مدیریت شهری فراهم کند. هدف این پژوهش شبیه‌سازی رشد کلانشهر تبریز با استفاده از الگوریتم‌های یادگیری ماشین متداول در مدل‌سازی تغییرات کاربری بویژه زمین‌های شهری و تلفیق آن با مدل سلول‌های خودکار بود. بنابراین برای شبیه‌سازی رشد شهری از نقشه‌های پتانسیل تغییر که براساس الگوریتم‌های جنگل تصادفی، ماشین بردار پشتیبان و شبکه عصبی پرسپترون چندلایه با پارامترهای بهینه استخراج گردید، استفاده شد. جهت جلوگیری از بیش‌برازش مدل‌ها و کسب نتایج خوشبینانه به‌دلیل همبستگی مکانی میان نمونه‌های آموزشی و آزمایشی، این الگوریتم‌ها با روش اعتبارسنجی متقابل مکانی ارزیابی شدند. نتایج حاصل نشان داد الگوریتم جنگل تصادفی با مقدار AUC معادل ۰/۹۲۲۸ از دقت و تعمیم‌پذیری بالاتری نسبت به سایر الگوریتم‌ها برخوردار می‌باشد. همچنین دقت مدل سلول‌های خودکار مبتنی بر پتانسیل تغییر به‌دست آمده از الگوریتم‌های بیان شده، با استفاده از سنجه سازگاری مورد ارزیابی و مقایسه قرار گرفت. نتایج این سنجه نیز نشان داد نقشه‌های پتانسیل تغییر حاصل از الگوریتم جنگل تصادفی در مدل سلول‌های خودکار، عملکرد بهتری نسبت به الگوریتم‌های ماشین بردار پشتیبان و شبکه عصبی پرسپترون چندلایه در پیش‌بینی تغییرات منطقه‌های غیر شهری به شهری در هر دو دوره کالیبراسیون (۱۳۷۶ - ۱۳۸۵) و اعتبارسنجی (۱۳۹۴ - ۱۳۸۵) داشته است. مدل CA مبتنی بر الگوریتم جنگل تصادفی در دوره اعتبارسنجی، ۱۳۷۶۱۰۰۰ مترمربع از زمین‌های رشد یافته شهری را بویژه در منطقه‌های داخلی شهر درست پیش‌بینی کرده است. تفاوت اصلی الگوریتم جنگل تصادفی نسبت به سایر الگوریتم‌ها به‌کارگیری روش یادگیری جمعی در برآورد

- ¹² Digital Elevation Model
¹³ United States Geological Survey
¹⁴ Landsat
¹⁵ Multilayer Perceptron Neural Network
¹⁶ Radial Basis Function
¹⁷ Ensemble
¹⁸ Classification and Regression Tree
¹⁹ Bootstrap
²⁰ Number of Trees
²¹ Mtry
²² Cross-Validation
²³ Overfitting
²⁴ Random search
²⁵ Grid search
²⁶ Moore
²⁷ Receiver Operating Characteristics
²⁸ True Positive Rate
²⁹ False Positive Rate
³⁰ Area Under the Curve
³¹ Figure of Merit
³² Local
³³ Confusion matrix

Aburas, M.M., Ho, Y.M., Ramli, M.F. and Ash'aari, Z.H., 2017. Improving the capability of an integrated CA-Markov model to simulate spatio-temporal urban growth trends using an Analytical Hierarchy Process and Frequency Ratio. *International Journal of Applied Earth Observation and Geoinformation*. 59, 65-78.

Ana, D., Nikolik, K. and Curfs, L., 2004. Probabilistic SVM outputs for pattern recognition using analytical geometry. *Neurocomputing*. 62, 293-303.

Asghari, A., 2015. Modeling Urban Development Using Cellular Automate and Ant Colony Optimization Algorithm Case Study: Tehran. MS.c. Thesis. Shahid Beheshti University, Thran, Iran.

Breiman, L., 2001. Random forests. *Machine Learning*, Springer. 45, 5-32.

Bergstra, J. and Bengio, Y., 2012. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*. 13, 281-305.

Bischi, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Studerus, E., Casalicchio, G. and Jones,

مقادیر خروجی است که می‌تواند اصلی‌ترین دلیل برتری این الگوریتم نسبت به دو الگوریتم مطرح دیگر مورد استفاده در پژوهش حاضر یعنی شبکه عصبی مصنوعی و ماشین بردار پشتیبان باشد.

پی‌نوشت‌ها

- ¹ Cellular Automata
² Transition Rules
³ Artificial Neural Network
⁴ Support Vector Machine
⁵ Random Forest
⁶ Shenzhen
⁷ Feature Space
⁸ Hyperplane
⁹ Wallonia
¹⁰ Harare
¹¹ Spatial Cross-Validation

منابع

Z.M., 2016. mlr: Machine Learning in R. *The Journal of Machine Learning Research*. 17, 5938-5942.

Clarke, K.C., Hoppen, S. and Gaydos, L., 1997. A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Environment and Planning B: Planning and Design*. 24, 247-261.

Duarte, E. and Wainer, J., 2017. Empirical comparison of cross-validation and internal metrics for tuning SVM hyperparameters. *Pattern Recognition Letters*. 88, 6-11.

Feng, Y., Liu, M., Chen, L. and Liu, Y., 2016. Simulation of dynamic urban growth with partial least squares regression-based cellular automata in a GIS environment. *International Journal of Geo-Information*. 5, 243.

Gislason, P.O., Benediktsson, J.R. and Sveinsson, J.A., 2006. Random forests for land cover classification. *Pattern Recognition Letters*. 27, 294-300.

Ghasemi Esfahan, A., 2013. Investigation of

- stochastic forest method to improve urban land cover classification using satellite images. MS.c. Thesis. Khaje Nasireddin Toosi University, Tehran, Iran.
- He, J., Xia, L., Yao, Y., Ye, H. and Jinbao, Z., 2018. Mining transition rules of cellular automata for simulating urban expansion by using the deep learning techniques. *International Journal of Geographical Information Science*. 32, 1362-3087.
- Hosseinali, F., Alesheikh, A.A. and Nourian, F., 2013. Agent-based modeling of urban land-use development, case study: Simulating future scenarios of Qazvin city. *Cities*. 31, 105-113.
- Hutter, F., Hoos, H.H. and Leyton-Brown, K., 2011. Sequential model-based optimization for general algorithm configuration. In *Lecture Notes in Computer Science*. 6683, 507-523.
- Huang, B. and Boutros, C., 2016. The parameter sensitivity of random forests, *BMC Bioinformatics*. 17, 331.
- James, G., Witten, D., Hastie, T. and Tibshirani, R., 2013. *An introduction to statistical learning*. Springer, New York, USA.
- Jat, M.K., Choudhary, M. and Saxena, A., 2017. Application of geo-spatial techniques and cellular automata for modelling urban growth of a heterogeneous urban fringe. *The Egyptian Journal of Remote Sensing and Space Science*. 20(2), 223-241.
- Javadi, Y., 2008. Modeling Land cover changes using Cellular Automata in GIS environment. MS.c. Thesis. University of Tehran, Tehran, Iran.
- Kamusoko, C. and Gamba, J., 2015. Simulating Urban Growth Using a Random Forest-Cellular Automata (RF-CA) Model. *ISPRS International Journal of Geo-Information*. 4, 447- 470.
- Kiavarz Moghaddam, H. and Samadzadegan, F., 2009. Urban simulation using neural networks and cellular automata for land use planning. In: M. Schrenk, et al., eds. *Proceeding of REAL CORP, Tagungsband*, pp. 571-577.
- Li, X. and Gong, P., 2016. Urban growth models: Progress and perspective. *Science Bulletin*. 61, 1637-1650.
- Liu, X. Li, X. and Shi, X., 2008. Simulating complex urban development using kernel-based non-linear cellular automata. *Ecological Modelling*. 211, 169-181.
- Liu, X. Li, X. and Chen, Y., 2010. A new landscape index for quantifying urban expansion using multi-temporal remotely sensed data. *Landscape Ecology*. 25, 671- 682.
- Liu, Y. Feng, Y. and Pontius, R., 2014. Spatially-explicit simulation of urban growth through self-adaptive genetic algorithm and cellular automata modelling. *Land*. 3, 719-738.
- Liang, X. Liu, X. Li, D. Zhao, H. and Chen, G., 2018. Urban growth simulation by incorporating planning policies into a CA-based future land-use simulation model. *International Journal of Geographical Information Science*. 32, 2294-2316.
- Lovelace, R. Nowosad, J. and Muenchow, J., 2019. *Geocomputation with R: Statistical learning*. CRC Press.
- Mundia, C. N. and Aniya, M., 2007. Modeling urban growth of Nairobi city using cellular automata and geographical information systems. *Geographical Review of Japan*. 80, 777-788.
- Maleki D., 2010. Modeling Urban Development Using Cellular Automation Method. M.Sc. Thesis. Khaje Nasireddin Toosi University of Technology,

Tehran, Iran.

Mirbagheri, B. and Alimohammadi, A., 2018. Integration of local and global support vector machines to improve urban growth modelling. *ISPRS International Journal of Geo-Information*. 7, 347.

Moosavi, M., 2011. An introduction to environmental challenges of life in slum settlements of Tabriz. 2th International Conference on Humanities. Historical and Social Sciences, 26th -28th February, Singapore. P.17.

Mustafa, A., Rienow, A., Saadi, I., Cools, M. and Teller, J., 2018. Comparing support vector machines with logistic regression for calibrating cellular automata land use change models. *European Journal of Remote Sensing*. 51, 391- 401.

Platt, J. C., 1999. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in Large Margin Classifiers*. 68, 61-74.

Pijanowski, B. C., Brown, D. G., Shellitoc, B. A. and Manikd, G. A., 2002. Using neural networks and GIS to forecast land use changes: a land transformation model. *Computers. Environment and urban systems*. 6, 553-575.

Pontius, G. R. and Malanson, J., 2005. Comparison of the Structure and Accuracy of Two Land Change Models. *International Journal of Geographical Information Science*. 19, 243-265.

Pontius G. R., Walker, R., Kumah, R., Arima, E., Aldrich, S., Caldad, M. and Vergara, D., 2007. Accuracy assessment for a simulation model of Amazonian deforestation. *Annals of the American Association of Geographers*. 97, 677-695.

Qian, Y., Xing, W., Guan, X., Yang, T. and Wu, H., 2020. Coupling cellular automata with area

partitioning and spatiotemporal. *Science of the Total Environment*. 722, 137738.

Rumelhart, D., Hinton, G. and Williams, R., 1986.. Learning representations by back-propagating errors. *Nature*. 323, 533-536.

Shafizadeh-Moghadam, H., Asghari, A., Tayyebi, A. and Taleai, M., 2017. Coupling machine learning, tree-based and statistical models with cellular automata to simulate urban growth. *Computers. Environment and Urban Systems*. 64, 297-308.

Tobler, W. R., 1970. A computer movie simulating urban growth in the Detroit region. *Economic geography*. 46, 234-240.

UNFPA, 2016. State of world population 2016. New York, United Nations Population Fund.

Widodo, A., Yang, B.S. and Han, T., 2007. Combination of independent component analysis and support vector machines for intelligent faults diagnosis of induction motors. *Expert systems with applications*. 32, 299-312.

White, R. and Engelen, G., 1997. Cellular automata as the basis of integrated dynamic regional modeling, *Environment and Planning B: Planning & Design*. 24, 235- 246.

Wu, F. and Webster, C. J., 1998. Simulation of land development through the integration of cellular automata and multicriteria evaluation. *Environment and Planning B: Urban Analytics and City Science*. 25, 103-126.

Wu, F. and Martin, D., 2002. Urban expansion simulation of Southeast England using population surface modelling and cellular automata. *Environment and Planning*. 34, 1855-1876.

Wu, T. F., Lin, C. J. and Weng, R. C., 2004.

Probability estimates for multi-class classification by pairwise coupling. *Journal of Machine Learning Research*. 5, 975-1005.

Yang, Q., Li, X. and Shi, X., 2008. Cellular automata for simulating land use changes based on support vector machines. *Computers & Geosciences*. 34, 592-602.

Yao, Y., Li, X., Liu, X., Liu, P. and Liang, Z., 2017. Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model. *International Journal of Geographical Information Science*. 31, 825-848.

Yeh, A.G., Li, X. and Xia, C., 2021. Cellular automata modeling for urban and regional planning. *Urban Informatics*. 45,865-883.

Zhang, Y., Liu, X., Chen, G. and Hu, G., 2020. Simulation of urban expansion based on cellular automata and maximum entropy model. *Science China Earth Sciences*. 63, 701-712.





Environmental Sciences Vol.19 / No.4 / Winter 2022

183-204
Original Article

Urban growth simulation using cellular automata model and machine learning algorithms (case study: Tabriz metropolis)

Omid Ashkriz, Babak Mirbagheri*, Ali Akbar Matkan and Alireza Shakiba

Center for Remote Sensing and GIS Research, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran

Received: 2020.11.23 Accepted: 2021.09.05

Ashkriz, O., Mirbagheri, B., Matkan, A.A. and Shakiba, A.R., 2022. Urban growth simulation using cellular automata model and machine learning algorithms (case study: Tabriz metropolis). *Environmental Sciences*. 19(4): 183-204.

Introduction: Urban growth has accelerated in recent decades, therefore, predicting the future growth pattern of the city is very important to prevent environmental, economic, and social problems. The city of Tabriz has also experienced rapid growth of urban lands due to significant demographic changes, which requires accurate simulation of urban growth to prevent negative environmental and economic consequences. The purpose of this study was to evaluate the performance accuracy of the proposed machine learning algorithms by spatial cross-validation method in combination with the cellular automata model to simulate urban growth.

Material and methods: In this study, to analyze urban land-use changes, Landsat satellite images related to the years 1997, 2006, and 2015 were classified using the support vector machine algorithm. In the next step, change potential maps of non-urban to urban areas were produced using random forest algorithms, support vector machine, and multilayer perceptron neural network for two periods of calibration (1997 and 2006) and validation (2006 and 2015) based on distance from the main roads, distance from the city center, distance from built-up areas, distance from the rivers and railways, as well as slope, elevation, and two-class (agricultural/barren) land use layer as effective factors in the growth of the city. Finally, using the cellular automata model, the growth simulation of Tabriz city based on land use and change potential maps obtained from machine learning algorithms for the mentioned periods was performed. To prevent over-fitting of algorithms to training samples and to obtain optimistic results, in the process of extracting optimal parameters of machine learning algorithms, the spatial cross-validation method was used to reduce the spatial correlation between training and test data.

* Corresponding Author: *Email Address.* b_mirbagheri@sbu.ac.ir

Results and discussion: The results showed that the random forest algorithm with the area under the ROC curve of 0.9228 compared to the support vector machine and multilayer perceptron neural network algorithms with 0.8951 and 0.8726, respectively, had a better performance in estimating the change potential of non-urban to urban areas. Furthermore, in comparison with others, the random forest also clearly showed local variations in potential change. Finally, the growth of Tabriz city was simulated using the cellular automata model based on the obtained change potential maps. Comparison of the prediction map in the validation period with the current situation of urban areas in 2015 showed that the accuracy of an urban growth simulation model based on random forest with a Figure of Merit index of 0.3569 compared to models based on support vector machine and artificial neural network was more accurate in allocating non-urban to urban lands with 0.3496 and 0.3434, respectively.

Conclusion: As machine learning algorithms such as artificial neural networks, support vector machines, and random forest are capable of solving non-linear problems, using them is strongly recommended for urban growth simulation. Also, among the algorithms used in this research, the random forest algorithm based on ensemble learning has a higher advantage than the two-support vector machine and the artificial neural network algorithms.

Keywords: Urban growth simulation, Cellular automata, Random Forest, Support vector machine, Artificial neural network.