

مدل مفهومی الگویابی پارسل های مناسب مسکونی در شهرهای جدید با استفاده از داده کاوی مکان محور (مطالعه موردی شهر جدید پردیس)

مرضیه بوربور اژدری، مهدیه بوربور اژدری

کارشناسی ارشد مدیریت فناوری اطلاعات، دانشگاه خوارزمی

چکیده

در پی سیاست ایجاد شهرهای جدید با هدف مهار رشد بی رویه جمعیت و کنترل مهاجرت به شهرهای بزرگ، بر اساس مصوبه سال ۱۳۶۴ هیأت وزیران شهرهای جدید در سطح کشور تاسیس گردیدند. تا کنون مهندسان و مشاوران وزارت راه و شهرسازی؛ جهت تهیه طرح های جامع و تفصیلی شهرهای جدید هیچ گونه الگوی مبتنی بر تکنیک های فناوری اطلاعات در اختیار نداشته و تنها بر اساس تجربه و اصول مهندسی اقدام به طراحی طرح های فوق الذکر نموده اند. در این پژوهش تلاش شده با بهره برداری از اطلاعات ارزشمند بانک اطلاعات مکانی (GIS) شرکت عمران شهر جدید پردیس و همچنین به کمک فرآیند داده کاوی، مدلی مفهومی جهت کشف الگوهای مناسب در طراحی پارسل ها با کاربری مسکونی ارائه گردد. در این مدل به علت نوع بانک اطلاعاتی که دارای داده های مکانی و توصیفی است، از "داده کاوی جغرافیایی" که نوعی خاص از داده کاوی فضایی می باشد، استفاده شده است. الگو طراحی شده برای معماران، مهندسان و مشاوران شهرسازی در ارائه طرح های جامع و تفصیلی شهرهای جدید مفید بوده و می تواند به افزایش استقبال شهروندان برای سرمایه گذاری و سکونت در سطح این شهرها بواسطه تامین بودجه شهرهای جدید از طریق واگذاری پارسل ها با کاربری های مختلف به متقاضیان حقیقی و حقوقی کمک کند. همچنین اگر بتوان به مهندسان طرح های جامع و تفصیلی، الگویی از پارسل هایی که با اقبال عمومی مواجه شده اند ارائه کرد. آنها می توانند از این الگو در طراحی سایر شهرها، به منظور افزایش فروش و بودجه شرکت های عمران اقدام نمایند.

واژه های کلیدی: شهر جدید پردیس، بانک GIS، پارسل، داده کاوی فضایی، قوانین همبستگی یا انجمنی، دسته بندی، درخت

تصمیم

مقدمه :

الگوی شهرهای جدید به عنوان پایه و اساسی برای سازماندهی و پالایش شهرهای بزرگ انتخاب شده است. این شهرها درمدهای مختلف اقماری، مستقل، بازسازی شده، پایدار، اداری و سیاسی در اروپا، آمریکا، استرالیا، آسیا و آفریقا برنامه‌ریزی و ساخته شده‌اند. هدف از ایجاد شهرهای جدید را می‌توان جذب سرریز جمعیتی شهرها، کاهش بار فعالیت‌های اقتصادی مادرشهر، توسعه ناحیه ای و استفاده از منابع طبیعی و همچنین توزیع بهینه جمعیت عنوان کرد. بدین ترتیب در واقع این شهرها برای کمک به حل مشکلات اقتصادی و اجتماعی و محیطی شهرهای بزرگ به وجود آمدند. شهرهای جدید پاسخی به تمرکز بیش از حد جمعیت و فعالیت‌های اقتصادی درون شهرهای بزرگ بوده‌اند. در ایران، رشد سریع جمعیت شهری در شهرهای بزرگ کشور و لزوم هدایت سرریز جمعیت و فعالیت‌های اقتصادی در آینده از شهرهای بزرگ مانند تهران، اصفهان، تبریز، شیراز و مشهد به سوی شهرهای جدید، ایجاد آنها را ضروری ساخته است. به عبارت دیگر، به دنبال مسائل اقتصادی اجتماعی و کالبدی کلان شهرهای ایران (به ویژه کلان شهر تهران)، راه حل ایجاد شهرهای جدید همانند بسیاری از کشورهای جهان مطرح گردید. چهار شهر جدید طرف کلانشهر تهران که عبارت‌اند از: پرنده، پردیس، اندیشه و هشتگرد نیز با چنین اهدافی طرح ریزی و احداث شده‌اند. توزیع مناسب و برنامه ریزی شده جمعیت در ناحیه شهری تهران از طریق هدایت سرریز جمعیت کلانشهر تهران به شهرهای جدید از جمله سیاست‌های عمده ایجاد و توسعه شهرهای جدید اطراف تهران بوده است. از سوی دیگر، میزان موفقیت شهرهای جدید و از جمله شهرهای جدید اطراف تهران به عوامل زیادی بستگی دارد؛ یکی از این عوامل نوع طراحی پارسل‌ها با کاربری‌های مختلف در طرح‌های جامع و تفصیلی این شهرها می‌باشد. داشتن یک الگوی مناسب از پارسل‌ها با کاربری مسکونی می‌تواند در برنامه ریزی شهری و شهرسازی شهرهای جدید راهگشا باشد. (قرخلو، ۱۳۸۸)

در طی بیست سال اخیر از روش داده‌کاوی فضایی در طراحی پارسل‌های پروژه‌های جامع و تفصیلی شهرهای جدید استفاده نشده است. این پژوهش سعی دارد با استفاده از این روش، عملکرد شرکت‌های عمران شهرهای جدید را در جذب جمعیت و سرمایه‌گذاری در این شهرها بهبود بخشد.

جهت بررسی مسئله، شهر جدید پردیس به عنوان مطالعه موردی انتخاب شده است. این شهر براساس مصوبه بیست و سوم اسفند ۱۳۶۸ هیأت وزیران، به مساحت ۳۸۰۰ هکتار و ارتفاع متوسط ۱۸۰۰ متر از سطح آب‌های آزاد، در کیلومتر ۱۱۷ آزاد راه تهران پردیس، با جمعیت نهایی دویست هزار نفر به عنوان نزدیک‌ترین شهر جدید به کلان شهر تهران احداث گردید.

شرکت عمران شهر جدید پردیس به عنوان متولی ساخت این شهر از سال ۱۳۸۶ اقدام به طراحی و راه‌اندازی بانک GIS این شهر نمود و این پروژه در حال حاضر با همکاری فرمانداری، شهرداری و سایر ارگان‌های مربوطه این شهر در حال تکمیل اطلاعات می‌باشد (قرخلو، ۱۳۸۸). با استفاده از اطلاعات ارزشمند بانک اطلاعات مکانی شهر و به کمک فرآیند داده‌کاوی در این پژوهش مدل مفهومی از الگوهای کشف شده‌ی پارسل‌های مناسب مسکونی به منظور آماده‌سازی این پارسل‌ها جهت افزایش استقبال شهروندان برای سرمایه‌گذاری و سکونت در سطح شهر جدید ارائه شده است.

در بانک GIS شهر جدید پردیس داده‌های مکانی و توصیفی شهر به تفکیک فازها و محله‌ها و پارسل‌های مختلف وجود دارد که می‌تواند حاوی اطلاعات و الگوهای با ارزشی باشد ولیکن تاکنون هیچ‌گونه فرآیند کشف دانش بر روی این بانک صورت نگرفته و تنها به جمع‌آوری داده‌ها اکتفا شده است.

با رشد نمایی حجم داده‌ها در بانک‌های اطلاعاتی مکانی؛ تجزیه، تحلیل و تفسیر دستی مجموعه‌های داده‌ای، بسیار کند، گران و غیر عملی می‌گردد. در حالیکه ارزش واقعی چنین داده‌هایی در توانایی استخراج گزارش‌های مفید و شناسایی روند رویدادهای جالب و به‌طور خلاصه کشف دانش می‌باشد (مینایی، شکوهی، امیرحسینی و بهروز، ۱۳۸۶؛ بهروز مینایی بیگدلی، مریم نظری دوست، آذر ۱۳۹۱)

دانش داده کاوی با بهره گیری از علوم مختلف همچون آمار، ریاضیات و هوش مصنوعی به پیش بینی و کشف الگوهای پنهان در داده‌ها می پردازد. امروزه کاربرد داده کاوی در دنیا بطور فزاینده ای در حال گسترش است اما در ایران در حال معرفی است. از نمونه های کاربردهای داده کاوی در کشورمان می توان به پیش بینی هزینه در صنعت بیمه، پیش بینی وضعیت جوی و رخدادهای آن در هواشناسی، یافتن پولشویی در نظام بانکی، کاهش هزینه های تعمیرات نگهداری تجهیزات در صنعت خودرو، بررسی سبد خرید مشتریان در فروشگاه های زنجیره ای و کشف قوانین پنهان در آن و بررسی وضعیت شرکتهای بدهکار در سازمان تامین اجتماعی اشاره کرد (محمدیان، ۱۳۹۱).

در خصوص کاربرد داده کاوی در زمینه مدیریت شهری از داده کاوی فضایی استفاده می شود. کاوش داده ها روی داده هایی که دارای یک یا چند ویژگی مکانی، فضایی و جغرافیایی باشند داده کاوی فضایی نامیده می شود و خروجی آن دانش و معرفتی است که دارای خصوصیات فضایی و جغرافیایی مانند مکان، جهت، فاصله، شکل هندسی و مانند آن باشد. داده کاوی جغرافیایی نوعی از داده کاوی فضایی است که اشیا و یا خصیصه های فضای جغرافیایی را در نظر می گیرد. (اسلامی، ۱۳۸۶)

استفاده از تکنیک های پیشرفته داده کاوی مانند: پیش بینی، طبقه بندی، خوشه بندی و ... میتواند در طبقه بندی پارسل های مسکونی و یافتن الگوهای خاص و با ارزش در مورد پارسل های مناسب به ما کمک کند.

با توجه به توضیحات فوق، در این مقاله در ابتدا مبانی نظری و پیشینه پژوهش شامل مفاهیم داده کاوی، GIS و داده کاوی، داده کاوی فضایی، روش های داده کاوی و تحقیقات قبلی بیان می شود، در بخش دوم روش شناسی تحقیق مطابق استاندارد-CRISP DM در چارچوب درک کسب و کار، شناسایی و درک داده ها، آماده سازی داده ها، مدل سازی و ارزیابی تشریح می گردد. پس از بیان متدولوژی، مدل مفهومی الگویابی پارسل های مناسب ارائه و در بخش پایانی به جمع بندی و نتیجه گیری بحث پرداخته شده است.

مبانی نظری و پیشینه تحقیق

مفهوم داده کاوی

داده کاوی به بررسی و تجزیه و تحلیل مقادیر عظیمی از داده ها به منظور کشف الگوها و قوانین معنی دار اطلاق می شود. خصوصیات اصلی داده کاوی، کشف خودکار داده ها، پیشگویی نتایج محتمل، ایجاد اطلاعات قابل استفاده و تمرکز بر روی مجموعه داده ها و پایگاه داده های بسیار بزرگ است. داده کاوی قادر است پرسش هایی که از طریق پرس و جوهای ساده و تکنیک های گزارشگیری قابل دستیابی نیستند، پاسخ دهد. (سجادیان، ۱۳۸۹)

GIS و داده کاوی در آن

آنچه داده کاوی در سیستم GIS را از انواع دیگر داده کاوی متمایز می کند، وجود داده های مکانی در GIS است. علاوه بر این GIS به دلیل داشتن واسط کاربر قوی، می تواند به عنوان ابزاری برای بصری سازی دانش های کشف شده مورد استفاده قرار گیرد (اسلامی، ۱۳۸۶). سیستم GIS غیر از آنکه منبع اطلاعاتی است، حاوی مجموعه ای از شبکه ها از جمله شبکه معابر، شبکه برق، شبکه آب، شبکه مخابرات، شبکه فاضلاب و غیره می باشد و به دلیل وجود این شبکه، سیستم تحلیل شبکه نیز در GIS وجود دارد. ساختار داده ها در GIS به این صورت است که به علت حجم زیاد داده ها، اطلاعات معمولاً در چندین لایه ذخیره می شوند. در این ساختار از هر لایه به صورت مجزا استفاده می شود و سیستم GIS و امکان تغییر دادن اطلاعات هر لایه بدون آنکه تغییری در دیگر لایه را ایجاد کند را فراهم می کند. (سجادیان، ۱۳۸۹)

داده کاوی فضایی

داده کاوی فضایی یکی از شاخه های داده کاوی است که در مورد استخراج دانش نهفته از میان داده های فضایی بحث می کند. چون در سیستم هایی که با داده های فضایی کار می کنند (مثل سیستم های GIS و سیستم های سنجش از راه دور) حجم اطلاعات بسیار بالا است. داده های فضایی حاوی اطلاعات توپولوژیکی و یا مسافتی هستند و اغلب توسط مکانیزم های شاخص بندی سازماندهی شده و توسط متدهای دسترسی فضایی مورد دسترسی واقع می شوند (محمدی، ۱۳۸۹). خواص فضایی در پایگاه های داده ای به طور مستقیم ذخیره نمی شوند. بلکه با ادغام پایگاه داده رابطه ای با خواص فضایی این کار انجام می شود. این اطلاعات می توانند بوسیله سیستم های اطلاعات جغرافیایی نمایش داده شوند.

روش های داده کاوی

به منظور ایجاد شناخت بهتر از پایگاه داده های بسیار بزرگ و تحلیل آنها، می توان از روشهای یادگیری هدایت شده و یادگیری هدایت نشده استفاده کرد. روش یادگیری هدایت نشده، با فرآیندی خودکار و بی نظارت ساختار موجود در داده ها را آشکار میکند و روش یادگیری هدایت شده شامل دو فرآیند با عنوان فرآیند آموزش (یادگیری) و فرآیند ارزیابی می باشد (بهرز مینایی بیگدلی، مریم نظری دوست، آذر ۱۳۹۱). در روش طبقه بندی مجموعه داده اولیه به دو مجموعه داده با عنوان مجموعه داده های آموزشی و مجموعه داده های آزمایشی تقسیم می شود، با استفاده از مجموعه داده های آموزشی مدل ساخته می شود و از مجموعه داده های آزمایشی برای اعتبار سنجی و محاسبه دقت مدل ساخته شده استفاده می شود. انواع الگوریتم های طبقه بندی عبارتند از: روش های مبتنی بر درخت تصمیم، روشهای مبتنی بر قانون، استدلال مبتنی بر حافظه، شبکه های عصبی، روش های مبتنی بر نظریه بیز، ماشین های بردار پشتیبان (دکتر محمد صنیعی آباد، مهندس سینا محمودی، مهندس محدثه طاهرپور، ۱۳۹۱).

الگوریتم درخت تصمیم:

در الگوریتم های دسته بندی مبتنی بر درخت تصمیم دانش خروجی به صورت یک درخت از حالات مختلف مقادیر ویژگی ها ارائه میشود. نمایش دانش به شکل درخت سبب شده است که دسته بندی های مبتنی بر درخت تصمیم کاملاً قابل تفسیر باشند. در حالت کلی درخت تصمیم رسم شده برای یک مجموعه داده آموزشی، واحد و یکتا نیست. بر اساس یک مجموعه داده، درخت های تصمیم مختلفی می توان بدست آورد. در میان ویژگی های موجود ابتدا آنها که خاصیت جدا کنندگی بیشتری دارند انتخاب می کنیم تا اطلاعات بیشتری از داده ها برای ما فراهم کنند (دکتر محمد صنیعی آباد، مهندس سینا محمودی، مهندس محدثه طاهرپور، ۱۳۹۱).

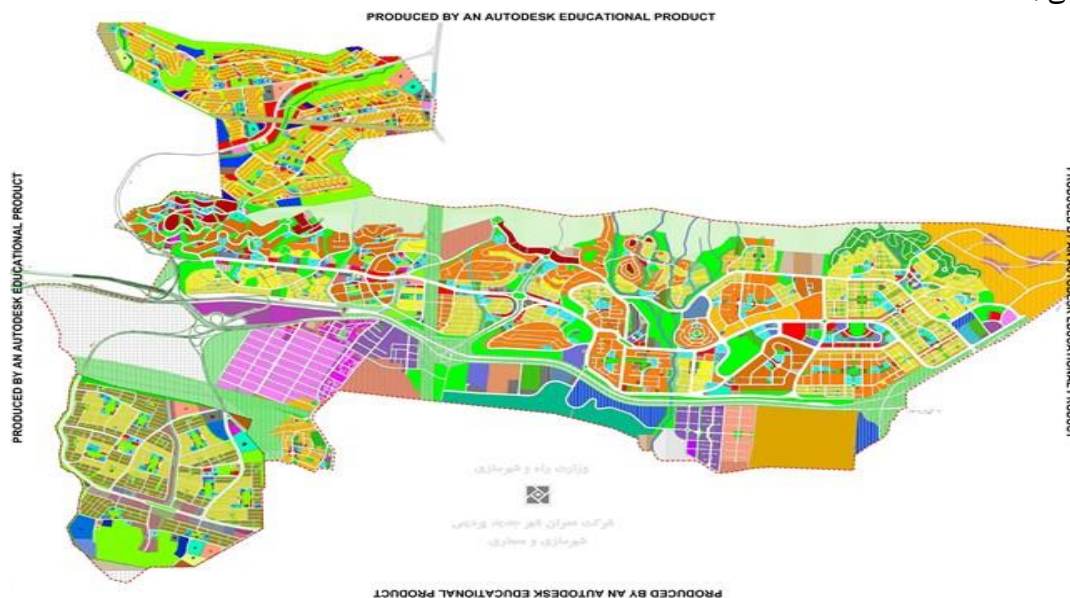
قوانین وابستگی:

یکی از روش های یادگیری هدایت نشده و بسیار مهم در داده کاوی می باشد. با استفاده از این تکنیک می توان روابط و وابستگی های جالبی در مجموعه داده پیدا کرد. استخراج قوانین وابستگی، زمانی انجام می شود که داده ها توصیف کننده وقایعی باشند که در یک زمان (یا زمان نزدیک به یکدیگر) رخ میدهند (بهرز مینایی بیگدلی، مریم نظری دوست، آذر ۱۳۹۱).

روش تحقیق

بر مبنای استاندارد CRISP-DM، فرآیند داده کاوی یک فرآیند چرخشی بوده که در شش فاز اصلی شامل درک مسأله و کار، شناسایی و درک داده ها، آماده سازی داده ها، مدل سازی، ارزیابی و گسترش اجرا می شود.

- تعریف مسئله/درک مسأله و کار: بر اساس اطلاعات بانک GIS پردیس؛ این شهر دارای ۱۱ فاز، ۴۴ محله و ۱۰۶۶۹ پارسل است. از این میان فاز های ۵، ۸، ۹، ۱۱ این شهر به علت اجرای طرح ملی مسکن مهر و فاز های ۱۰، ۷، ۶ به علت شرایط خاص از محدوده این پژوهش خارج شده و تمرکز این پژوهش بر روی فازهای ۳، ۲، ۱ و ۴ می باشد. در این میان؛ پارسل ها با کاربری مسکونی واقع در فاز ۲ محله ۲ دارای بالاترین ارزش ریالی بوده و با بیشترین استقبال شهروندان مواجه شده است. هدف ما تهیه الگویی از این فاز به کمک فرآیند داده کاوی و جهت پیاده سازی در فاز های ۳، ۱ و ۴ به منظور افزایش استقبال مردمی جهت سرمایه گذاری در این فازها و در نتیجه افزایش بودجه شرکت عمران شهر جدید پردیس می باشد.



- شناسایی و درک داده ها: از سال ۱۳۸۶ شرکت مادر تخصصی عمران شهرهای جدید اقدام به تهیه بانک جامع GIS شهرهای جدید کشور نمود. این بانک حاوی داده ها و اطلاعات با ارزشی در خصوص ۱- طرح و وضع موجود شبکه مخابرات، برق، آب و فاضلاب، گاز، حمل و نقل، معابر و فضای سبز ۲- طرح جامع، تفصیلی و آماده سازی شهر ۳- امور حقوقی، واگذاری، املاک، معماری و شهرسازی میباشد. مجموعه داده مورد استفاده در این پژوهش، تعداد ۱۰۶۶۹ (رکورد) از جدول قطعه بانک GIS میباشد. در جدول قطعه این بانک فیلد قیمت مربوط به هر قطعه وجود ندارد و از آنجا که در این پژوهش مبنای اولیه در طبقه بندی پارسل ها به عنوان پارسل مناسب یا نامناسب قیمت آن پارسل می باشد. (فیلد قیمت متغیر هدف می باشد). لذا جدولی جدید به نام "قیمت قطعه" به بانک اضافه شد که شامل فیلد های شماره قطعه، شماره قرارداد واگذاری، مساحت قطعه، قیمت هر متر مربع و قیمت کل قطعه میباشد. Forien Key در این جدول شماره قرارداد واگذاری میباشد و از طریق همین کلید خارجی دو جدول قطعه و قیمت قطعه به یکدیگر Join شدند. و برای محاسبه فیلد قیمت کل هر قطعه باید مساحت کل قطعه در قیمت هر متر مربع ضرب شود.
- آماده سازی داده ها: به منظور آماده سازی داده ها بخش هایی از بانک مانند شبکه های آب، برق، مخابرات، حمل و نقل، فاضلاب کنار گذاشته شده و بر روی بخش های شهرسازی و معماری، واگذاری، املاک و شبکه معابر متمرکز شدیم و با در نظر گرفتن شاخص های زیر جهت آماده سازی از نرم افزار SPSS Modeller استفاده می کنیم.
- ۱- Data Cleansing: عملیاتی که به برطرف شدن مشکل کیفیت داده ها می انجامد، پاک سازی داده نامیده میشود.

۱-۱- Noise: یک مقدار اپسیلون است که با داده اصلی جمع و یا تفریق میشود و موجب می گردد که به اصل داده دسترسی نداشته باشیم.

۱-۲- Duplicate Data: رکوردهایی که بار اطلاعاتی جدیدی ندارند و اطلاعات تکراری زیادی در آنها وجود دارد.

۱-۳- Out lier ها: یا همان نقاط پرت؛ نمونه رکورد هایی هستند که مقادیر ویژگی های آنها نسبت به سایر رکورد ها بسیار متفاوت است و موجب می شوند که الگوریتم یادگیری نتواند مدل دقیقی را برای تبیین دانش نهفته در داده بیابد. با استفاده از تکنیک خوشه بندی می توان داده پرت را به عنوان خوشه های تک عضوی در نظر گرفت و حذف نمود.

۱-۴- Missing Value ها: بعضی از فیلدهای مربوط به رکورد های مختلف دارای مقدار Null هستند. از چهار روش حذف کردن، تخمین زدن، نادیده گرفتن و جایگزین کردن میتوان با مشکل مقادیر از دست رفته برخورد کرد. (دکتر محمد صنیعی آباده، مهندس سینا محمودی، مهندس محدثه طاهرپور، ۱۳۹۱)

۲- انتخاب فیلد: بر اساس نظر خبره فیلد های شماره قطعه، کاربری، مساحت، تراکم، وضعیت واگذاری، وضعیت ساخت و ساز، شماره فاز، شماره محله، وضعیت آماده سازی، وضعیت پروانه ساخت، شماره قرارداد واگذاری، قیمت هر متر مربع و قیمت کل قطعه میباشد. (دکتر محمد صنیعی آباده، مهندس سینا محمودی، مهندس محدثه طاهرپور، ۱۳۹۱)

۳- انتخاب رکورد: از بین مجموع رکورد ها تنها رکورد هایی که فیلد کاربری در آنها دارای مقدار مسکونی (ویلائی/ آپارتمانی) می باشد انتخاب می گردند.

۴- ساخت فیلد هدف: فیلد هدف (میانگین/میان/چارک اول/چارک سوم/چارک چهارم) قیمت بازار هر قطعه بر اساس استقبال مردمی می باشد.

اینکه از میانگین یا میان یا چارک اول یا چارک سوم و یا چارک چهارم جهت تعیین معیار مناسب یا نامناسب بودن استفاده کنیم بر اساس نظر خبره می باشد. در این پژوهش برای شروع کار ما میان قیمت بازار هر قطعه را در نظر میگیریم. یعنی قیمت هر قطعه مسکونی با میان قیمت تمام قطعات مسکونی که در شعاع یک کیلومتری آن قطعه قرار دارند مقایسه میشود. چنان چه قیمت قطعه مورد نظر از میان (۵۰٪) قیمت بازار بیشتر باشد آن را پارسل مناسب مینامیم و چنانچه از میان (۵۰٪) قیمت بازار پایین تر باشد آن را پارسل نامناسب مینامیم.

- مدلسازی: در این مدل مفهومی جهت مدل سازی برای یافتن الگوهای مناسب پارسل ها، هر دو روش یادگیری هدایت شده و هدایت نشده را به کار برده و بدین منظور از دو تکنیک اصلی طبقه بندی و قوانین وابستگی (پیوند یا انجمنی) استفاده خواهیم کرد.

- ارزیابی: به منظور ارزیابی سه فاکتور؛ Support (میزان پشتیبانی مدل)، Confident (میزان اطمینان مدل)، Lift (میزان بهبود مدل) را مورد بررسی قرار خواهیم داد. (دکتر محمد صنیعی آباده، مهندس سینا محمودی، مهندس محدثه طاهرپور، ۱۳۹۱)

مدل مفهومی

مدل سازی :

۱- از آنجا که متغیر هدفی عنی پارسل ها با کاربری مسکونی دارای دو توزیع گسسته مناسب و نامناسب است. بنابراین از روش طبقه بندی (classification) که روش با ناظر (Supervised) است برای داده کاوی استفاده می کنیم.

۲- از طرفی چون به دنبال کشف الگوهای مربوط به مناسب بودن و نامناسب بودن هستیم؛ پژوهش ما در قالب توصیف مفهومی (Conceptual descriptions) قرار میگیرد.

از موارد ۱ و ۲ نتیجه میگیریم که باید از الگوریتم درخت تصمیم استفاده نماییم زیرا در مجموعه الگوریتم های روش طبقه بندی قرار دارد و دانش خروجی آن کاملاً قابل تفسیر و در قالب قانون می باشد.

تمام الگوریتم های درخت تصمیم برای شناسایی بهترین شکست از یک مکانیزم حریصانه استفاده میکنند که بر اساس این مکانیزم شکستی که توزیع دسته ها در گره های حاصل از آن همگن باشد، نسبت به سایر شکست ها بهتر است. در واقع گره همگن گره ای است که کمترین میزان ناخالصی را داشته باشد. در اینجا ما از سه روش Entropy, Gini Index و کایدو برای محاسبه ناخالصی گره استفاده میکنیم.

۳- به صورت موازی با گام دوم از روش قوانین وابستگی که از جمله روش های توصیفی بدون ناظر محسوب می شود، برای کشف سایر الگوهایی که توسط درخت تصمیم دیده نشده اند استفاده می کنیم. خروجی مهم در این روش کاوش داده مجموعه ای از قوانین اگر و آنگاه می باشد. (دکتر محمد صنیعی آباده، مهندس سینا محمودی، مهندس محدثه طاهرپور، ۱۳۹۱)

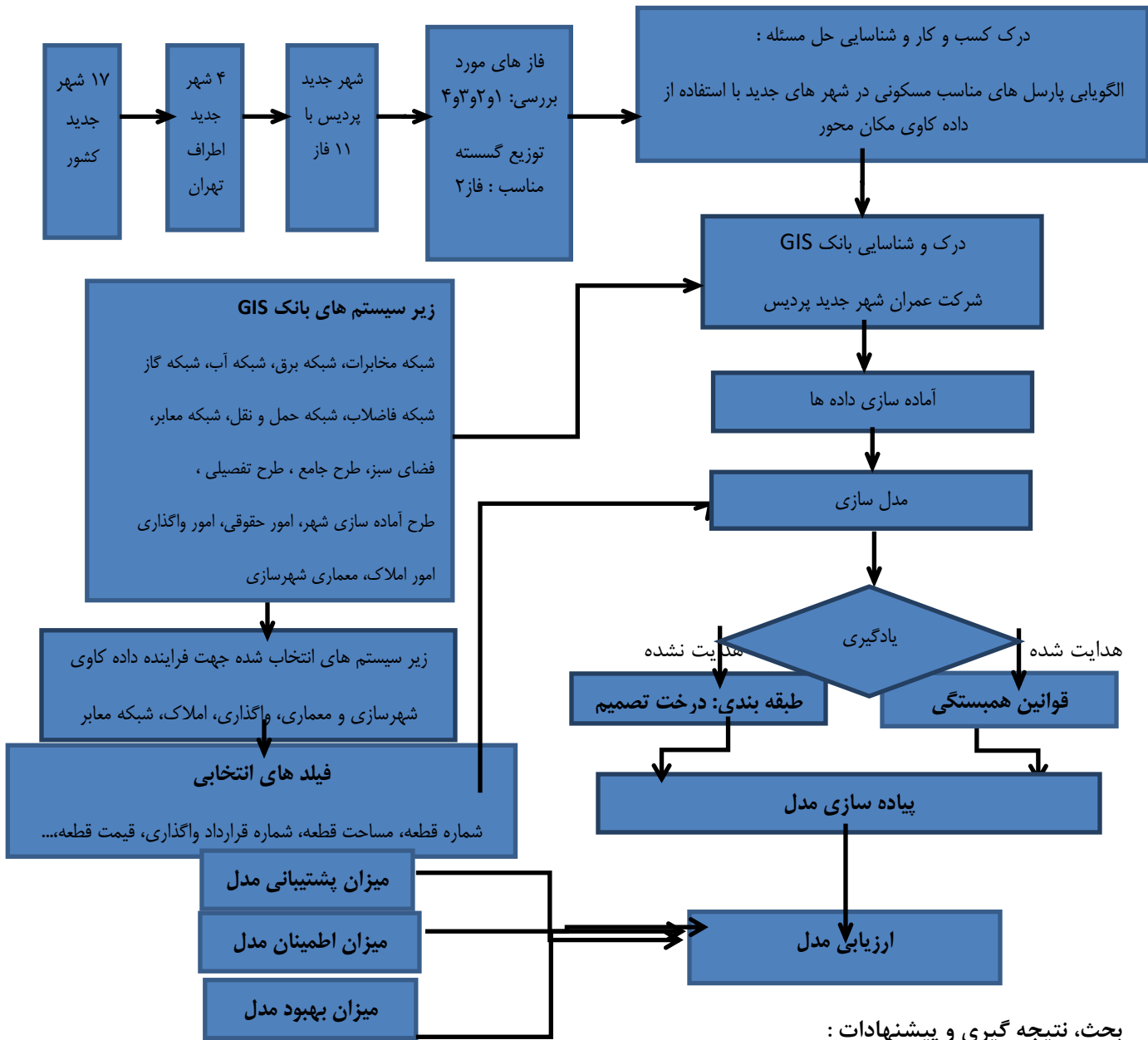
ارزیابی :

الگوها و قوانین ایجاد شده در یک مجموعه داده ممکن است بسیار زیاد باشد و همه این قانون ها لزوماً برای ما جذاب نباشد. از آنجا که جذابیت قوانین به عواملی همچون شخص استفاده کننده از قوانین و نیز حوزه ای ای که مجموعه داده مورد بررسی به آن تعلق دارد بستگی دارد، کار پیدا کردن قوانین جذاب، بدیهی به نظر نمیرسد. بنابراین مهم است که معیار های پذیرفته شده ای کیفیت قوانین مطرح نماییم:

- Support: میزان پشتیبانی قوانین از الگوها را مورد بررسی قرار میدهد.
- Confidence: میزان اطمینان از درست بودن قوانین را مشخص میکند.
- Lift: مقدار این شاخص از تقسیم Confidence قانون بر Support سمت راست قانون بدست می آید.

$$\text{Lift}(A \rightarrow B) = \frac{\text{Conf}(A \rightarrow B)}{\text{Sup}(B)}$$

این معیار میزان استقلال میان اشیاء A و B را بدست میاورد که بین صفر تا بی نهایت می باشد. (دکتر محمد صنیعی آباده، مهندس سینا محمودی، مهندس محدثه طاهرپور، ۱۳۹۱).



بحث، نتیجه گیری و پیشنهادات :

با توجه به سیاست گذاری های کلان دولتی جهت طراحی شهرهای جدید؛ به منظور افزایش میزان استقبال مردمی برای سرمایه گذاری و سکونت در این شهرها باید پارسل های شهری به گونه ای که بیشترین میزان رضایت شهروندان را فراهم نماید طراحی گردند. بدین منظور با استفاده از بانک اطلاعات مکانی (GIS) شهرهای جدید و فرآیند داده کاوی فضایی بر روی آن موفق به استخراج الگوهایی جهت پارسل های مناسب مسکونی شدیم. با توجه به اینکه تاکنون در زمینه طراحی پارسل های شهر سازی تنها از روش تجربی مهندسان شهر سازی استفاده شده است باید پس از بدست آوردن الگو های مربوط به پارسل های مناسب مسکونی به کمک داده کاوی مکانی، در فاز بعدی پژوهش با مهندسان شهرسازی (خبرگان) سیستم مصاحبه کرده و الگو های بدست آمده از این روش را با طرح های ایشان مقایسه نموده و میزان تطابق الگوهای بدست آمده از هوش مصنوعی با هوش انسانی را بررسی نماییم.

منابع و ماخذ :

۱. اسلامی، ح. ا. (۱۳۸۶). استخراج Landmarkها برای سیستم های هدایت مسیر عابر پیاده با استفاده از داده کاوی فضای چهارمین کنفرانس بین المللی مدیریت فناوری اطلاعات و ارتباطات، تهران، ندای اقتصاد بامداد (تاب)، ۲.
۲. پور، د. م. (۱۳۹۱). داده کاوی کاربردی. تهران: نیاز دانش.
۳. دوست، ب. م. (آذر ۱۳۹۱). داده کاوی پذیرفته شدگان مقطع کاردانی به روش پودمانی. ششمین کنفرانس داده کاوی ایران.
۴. سجادیان، م. م. س. (۱۳۸۹). مدیریت جامع سیلاب شهری توسط مدلسازیهای تلفیقی با دانش داده کاوی فضایی مبتنی بر چرخه دمینگ و بر اساس مهندسی ارزش. اولین کنفرانس ملی مدیریت سیلابهای شهری. تهران.
۵. قرخلو، م. (بهار ۱۳۸۸). ارزیابی عملکرد شهرهای جدید در جذب جمعیت کلان شهرها. پژوهش های جغرافیای انسانی (شماره ۶۷)، ۱۷-۲۵ صص.
۶. محمدی، س. ن. (۱۳۸۹). تجمیع فناوری های داده کاوی و کشف دانش با تحلیل داده های فضایی و سیستم اطلاعات جغرافیایی GIS. ششمین کنفرانس داده کاوی ایران.
۷. محمدیان، م. (۱۳۹۱). تاریخچه داده کاوی. <http://amargiran.com/Forum/forum-f15/topic-t268.html>
۸. مینایی، ش. ا. (۱۳۸۶). نقش داده کاوی GIS در مدیریت بحران. اولین همایش GIS شهری، دانشگاه شمال، ۱.