

تحلیل مصرف آب شهری (خانگی) شهرستان بابل با استفاده از روش‌های داده‌کاوی

مهرداد آهنگرکانی^۱

سید حسین خواسته^۲

تاریخ دریافت مقاله: ۹۷/۰۲/۱۵

تاریخ پذیرش مقاله: ۹۷/۰۷/۱۶

چکیده

مسئله کمبود آب در ایران با توجه به قرارگیری آن در منطقه خشک و کم آب خاورمیانه و روند نسبتاً سریع افزایش جمعیت شهری و مصرف بی‌رویه آب، هر روز ابعاد جدی‌تری به خود می‌گیرد. این مسئله و همچنین فقدان الگوی مشخص از نحوه و میزان مصرف آب در کشور، استفاده از سیستم‌های پردازش اطلاعات را در مدیریت منابع آب ضروری می‌نماید. کاوش قوانین وابستگی می‌تواند به عنوان یکی از روش‌های مطلوب جهت بهینه‌سازی و تحلیل عوامل مؤثر در میزان مصرف آب مورد استفاده قرار گیرد. در این مطالعه از کاوش قوانین وابستگی و الگوریتم درخت تصمیم به منظور تحلیل داده‌های موجود در زمینه مصرف آب شهری محله‌های شهرستان بابل استفاده شده است. پارامترهای مکانی و اجتماعی-اقتصادی مانند فاصله از رودخانه بابل‌رود، فاصله از راه‌های اصلی، ویلایی یا آپارتمانی بودن، مساحت فضای سبز خانگی، تعداد واحدهای مسکونی، تراکم جمعیت، درصد پیر یا جوان بودن، متوسط تعداد افراد هر خانواده و مساحت حیاط ساختمان انتخاب شدند. با استفاده از کاوش قوانین وابستگی به کشف ارتباط میان میزان مصرف آب و پارامترهای موردنظر پرداخته شده است. با بهره‌گیری از یک درخت تصمیم، عوامل مرتبط با مصرف، طبقه‌بندی و میزان مصرف آب بخشی از مشترکین به عنوان داده‌های تست پیش‌بینی شده است. به کمک تعیین نواحی با بیشترین میزان مصرف آب و الگوی مکانی توزیع مصرف آب خانوار، این قوانین ارزیابی شده‌اند. در این پژوهش پارامترهایی که رابطه عکس یا مستقیم با میزان مصرف آب خانوار دارند و همچنین پارامترهایی که اهمیت بیشتری نسبت به سایر عوامل دارند، مشخص شده‌اند. با بررسی میزان مصرف آب خانوار، محله‌های با بیشترین میزان مصرف که محله‌های ساحلی رودخانه بابل‌رود هستند، به عنوان محله‌های پرمصرف شناسایی شده‌اند. ساختمان‌ها در این محله‌ها، ویلایی و دارای تراکم جمعیت پایین‌تر و مساحت فضای سبز و حیاط بالاتری نسبت به سایر محله‌ها هستند.

واژه‌های کلیدی: داده‌کاوی، سیستم اطلاعات مکانی، کاوش قوانین وابستگی، درخت تصمیم، مصرف آب، شهرستان بابل.

۱- دانشجوی دکتری سامانه اطلاعات مکانی، دانشگاه صنعتی خواجه نصیرالدین طوسی (نویسنده مسئول) mahangar@mail.kntu.ac.ir

۲- استادیار مهندسی کامپیوتر، دانشکده کامپیوتر، دانشگاه صنعتی خواجه نصیرالدین طوسی khasteh@kntu.ac.ir

۱- مقدمه

شهری کمک و زمینه های لازم به منظور مدیریت تقاضای آب مسکونی را فراهم نماید (شهرستانی، ۱۳۹۳). مطالعات مختلفی در زمینه مصرف آب صورت گرفته است. فلاحی و همکاران در سال ۱۳۹۱، به بررسی تقاضای آب شهر نیشابور مربوط به ۲۶۶ خانوار در بازه زمانی ۱۳۸۲ تا ۱۳۸۶ پرداختند. نتایج به دست آمده از این مطالعه نشان دادند که اگر قیمت آب برای سال های آینده افزایش نیابد، با افزایش درآمد سرانه، مصرف سرانه آب خانوار افزایش می یابد. اما در صورت افزایش تعرفه ها در سال های آینده، مصرف سرانه آب خانوار در شبانه روز کاهش خواهد یافت (Fallahi, Ansari, & Moghaddas, 2012). همچنین صبوحی و همکاران در سال ۲۰۰۹، در مطالعه خود به این نتیجه رسیدند که افزایش قیمت آب در مواردی که اضافه مصرف بالاتر است، اثر بیشتری بر کاهش مصرف آب دارد (Sabouhi & NOUBAKHT, 2009). ابراهیمی و همکاران در سال ۱۳۸۰، روند رشد مصرف آب شامل مصرف خالص آب و مصرف سالیانه را طی سال های ۱۳۴۸ تا ۱۳۶۲ در شهر اصفهان مورد بررسی قرار دادند و به این نتیجه رسیدند که بهترین روش برای جلوگیری از بحران آب، روی آوردن به اقدامات مدیریتی جهت صرفه جویی در مصرف آب است که باید همسو با اهداف گسترده تر اقتصادی، اجتماعی و زیست محیطی عمل کند (ابراهیمی و نادری، ۱۳۸۰). جوادیان زاده و همکاران در سال ۱۳۸۸، در بررسی خود دریافتند که مشترکین شهر یزد در فصل تابستان بیشترین بهای آب را پرداخت نموده اند. با توجه به این امر که آب به عنوان یک کالای بدون جایگزین مطرح می باشد، افزایش یا کاهش قیمت آب در مقیاس کم بر میزان مصرف مردم اثر چندانی ندارد (جوادیان زاده، ۱۳۸۸). در سال ۱۳۹۱، آقاحسینی علی شیرازی با استفاده از روش فوریه در شهر بیرجند مشخص نمود که مقادیر مصرف پایه، مستقل از شرایط اقلیمی بوده و از روزهای تعطیل و عادی هفته و همچنین شماره روز هفته تأثیر می پذیرد. نتایج بیانگر این است که استفاده از سری فوریه در برآورد میزان مصرف روزانه کل، جواب خوبی نسبت به روش های دیگر نشان

آب مایه حیات و یکی از منابع بسیار مهم و محدود در جهان به شمار می آید. رشد جمعیت، روند شتابان شهرنشینی، افزایش مصرف و آلودگی منابع به همراه توزیع نابرابر آن در کره زمین، مشکلات فراوانی را برای بشر ایجاد کرده است. چالش های آبی در قاره آسیا نسبت به سایر مناطق جهان جدی تر و حادث تر است و کشورهای در حال توسعه نیز نسبت به کشورهای پیشرفته، آینده بحرانی تری در این زمینه پیش رو دارند. با توجه به موقعیت جغرافیایی ایران، اکثر مناطق کشور از نظر اقلیمی جزء مناطق خشک و نیمه خشک بوده و میانگین بارش در ایران حدود ۲۵۰ میلی متر است که این مقدار حدود یک سوم از میانگین بارش در جهان است. در سال های اخیر ایران با کمبود آب و مصرف زیاد منابع آبی توسط شهروندان مواجه بوده است. بنابراین توجه به الگوی مصرف آب شهری و در نظر گرفتن روابط بین پارامترهای جغرافیایی و مکانی، جمعیتی، اجتماعی، اقتصادی و غیره با میزان مصرف آب، یک مسئله مهم برای مدیریت مؤثر منابع آب محسوب می گردد (Statistical Center of Iran, ۲۰۱۱; وزیر، وکیل پور، & مرتضوی، ۱۳۹۵). ایران با وجود منابع کم آب شیرین، جزء پرمصرف ترین کشورهای جهان است و سرانه مصرف آب شهری در این کشور بیشتر از میانگین مصرف جهانی است (شهرستانی، ۱۳۹۳). اهمیت این مسئله و همچنین فقدان الگو و توصیف مشخص از نحوه و میزان مصرف آب در کشور، استفاده از سیستم های پردازش اطلاعات را در مدیریت منابع آب ضروری می نماید. لذا در جهت مدیریت مصرف آب شهری و ارائه تمهیدات گذر از بحران، لازم است اطلاعات لازم در رابطه با الگوی مصرف شهروندان بررسی شود. این اطلاعات شامل میزان مصرف هر یک از مصارف شهری به خصوص مصرف آب مسکونی، پراکندگی فضایی میزان مصرف آب و عوامل مؤثر بر الگوی مصرف می باشد. بر این اساس، بررسی الگوی مصرف آب شهری و استفاده اصولی از نتایج آن، می تواند به حل مسئله کمبود آب در تأمین آب

Malerba, 2003). ورسیچل در سال ۲۰۱۴، یک مورد مطالعاتی برای تحلیل بازدید از جاذبه‌های توریستی به وسیله کشف قوانین وابستگی مکانی را ارائه داده است.

در این مقاله بیان شده است که تمرکز روی جاذبه‌های توریستی مختلف با استفاده از قوانین وابستگی می‌تواند تقویت شود و می‌توان به اصلاح توزیع توریست‌ها در جاذبه‌های مختلف به جای تمرکز بر جاذبه‌های محدود پرداخت (Versichele et al., 2014).

از مدل درخت تصمیم و سیستم اطلاعات جغرافیایی برای پیش‌بینی توزیع خاک و سنگ بستر مستعد فرسایش در منطقه‌ای در لبنان استفاده شده است که این مدل برای مناطق دیگر مخصوصاً زمانی که اطلاعات خاک و سنگ منطقه محدود باشد قابل استفاده است (Kheir, Chorowicz, Abdallah, Dhont, 2008). همچنین از تکنیک درخت تصمیم جهت ارزیابی و پیش‌بینی طبقات خشکسالی در ترکیه استفاده شده است (Yurekli, Taghi Sattari, Anli, & Hinis, 2012) و کارایی مدل درخت تصمیم در پیش‌بینی مقدار بارش با استفاده از آمار و اطلاعات ایستگاه سینوپتیک یزد مورد بررسی قرار گرفته است (Dastourani, Habibipoor, Ekhtesasi, Talebi, & Mahjoobi, 2013). از مدل درخت تصمیم به همراه شبکه‌های عصبی مصنوعی جهت تحلیل وقوع سیلاب در حوزه تایوان (J.-C. Chen, Shu, Ning, & Chen, 2008) و جهت تجزیه و تحلیل تداوم خشکسالی‌ها در آنکارا استفاده شده است.

در این تحقیق از داده‌های ماهانه و سالانه بارش، دما، باد و رطوبت جهت ارزیابی وضعیت خشکسالی بهره گرفته شده است (Taghi Sattari, Anli, Apaydin, & Kodai, 2012).

در میان مطالعاتی که تاکنون صورت گرفته است، مطالعه میزان مصرف آب شهری با استفاده از الگوریتم‌های داده کاوی و با رویکرد مکانی انجام نشده است.

مسئله کمبود آب در ایران با توجه به قرارگیری آن در منطقه خشک و کم آب خاورمیانه و همچنین روند نسبتاً سریع افزایش جمعیت شهری و مصرف بی‌رویه آب، هرروز ابعاد جدی‌تری به خود می‌گیرد. این مسئله امروزه

می‌دهد (آفاحسینعلی، اکبریپور، ۱۳۹۰). محمدی در سال ۱۳۹۳، به بررسی تأثیر هدفمند یارانه‌ها بر الگوی مصرف و میزان آب شهر اردبیل پرداخته است. وجود رابطه معنادار میان هدفمندی یارانه‌ها و الگوی مصرف و کاهش مصرف آب در بخش‌های عمومی و صنعتی نشان می‌دهد که هدفمندی یارانه‌ها و افزایش قیمت آب به دلیل وسعت مصرف در این بخش‌ها باعث کاهش مصرف و اصلاح الگوی مصرف در این بخش‌ها شده است (محمدی، ۱۳۹۳). شمسایی در سال ۱۳۷۹، در مقاله‌ای به برآورد تابع تقاضای آب استان اصفهان پرداخته است که در آن برای کاهش یکسان مصرف آب در مناطق مختلف، افزایش قیمت‌های متفاوتی مورد نیاز خواهد بود. تثبیت قیمت آب و دادن یارانه به شدت از تأثیرگذاری دیگر عوامل بر تقاضا و به خصوص قیمت آب کاسته است (شمسایی، ۱۳۷۹). ملکی نسب و همکاران در سال ۱۳۸۶، استفاده از قطعات کاهنده مصرف شامل سرشیر و سردوش کم مصرف آب در سطح شهر کاشان را مورد ارزیابی قرار داده‌اند. بررسی تقاضا و تأمین آب شهر کاشان در سال‌های آینده نشان می‌دهد که استفاده از این قطعات در سطح شهر می‌تواند نیاز به اجرای طرح‌های آینده تأمین آب را تا ۶ سال به تأخیر اندازد و هزینه تهیه این قطعات برای مشترکین خانگی در مدت ۱۱ ماه با کاهش هزینه آب‌بها برگشت داده می‌شود، در حالی که عمر مفید قطعات ۵ سال می‌باشد (ملکی نسب، ابریشم‌چی و تجریشی، ۱۳۸۶). مسئله کاوش مجموعه آیت‌های متناوب برای اولین بار توسط آگروال در سال ۱۹۹۳ به صورت کاوش قوانین وابستگی در میان مجموعه آیت‌ها ارائه شد (Agrawal, Imieliński, & Swami, 1993).

یووم و کیم یک تکنیک فیلترینگ ترکیبی بر اساس الگوی رفتاری مشتریان و الگوهای مکانی در سایت‌های تجاری ارائه دادند (Kim & Yum, 2011).

برخی از قوانین وابستگی کشف شده می‌توانند نمایانگر دانش جدید در سطوح شهری باشند و در هدایت منابع برای بهبود امکانات به خصوص در نواحی با دسترسی ضعیف به حمل و نقل کمک کنند (Appice, Ceci, Lanza, Lisi, &

دوم و سوم به ترتیب منطقه مورد مطالعه در این تحقیق، داده‌های مورد استفاده و روند کلی تحقیق تشریح شده‌اند. روش‌های داده کاوی و آمار فضایی همچون کاوش قوانین وابستگی، درخت تصمیم و تحلیل زمین- آماری با استفاده از شاخص Moran's I در قالب مبانی نظری تحقیق در بخش چهارم توضیح داده شده‌اند. در بخش پنجم، نتایج تحقیق مورد بحث و تجزیه و تحلیل قرار گرفته و در نهایت، در قسمت پایانی، به جمع‌بندی و نتیجه‌گیری مطالب ارائه شده در فصول قبلی و پیشنهادات و راهکارهایی برای مطالعات آینده پرداخته شده است.

۲- منطقه مورد مطالعه

مسئله کمبود آب در شهرستان بابل با توجه به موقعیت جغرافیایی این شهرستان به عنوان یکی از شهرستان‌های کم‌آب در استان مازندران، مصرف بی‌رویه آب شهروندان و انجام فعالیت‌های پر مصرف از جمله کشاورزی، هر روز ابعاد جدی‌تری به خود می‌گیرد. از این رو، انتخاب آب این شهرستان جهت مطالعه ضروری به نظر می‌رسد. شهرستان بابل پرجمعیت‌ترین شهرستان استان مازندران می‌باشد. مرکز این شهرستان شهر بابل می‌باشد که به شهر بهارنارنج معروف است.

ارتفاع در این شهرستان از شمال تا ده کیلومتری جنوب آن، از سطح دریاهای آزاد پایین‌تر است. ارتفاع شهر بابل ۲ متر پایین‌تر از سطح دریاهای آزاد است. در سرشماری سال ۱۳۹۵ جمعیت شهرستان بابل ۵۳۱۹۳۰ نفر بوده و مساحت این شهرستان حدود ۱۵۷۸/۱ کیلومتر مربع است. این شهرستان در بخش شمالی ایران، در طول جغرافیایی ۵۲ درجه و ۴۰ دقیقه و ۲۰ ثانیه شرقی و عرض جغرافیایی ۳۶ درجه و ۳۳ دقیقه و ۱۰ ثانیه شمالی واقع است (نگاره ۱). شهرستان بابل متشکل از ۶ بخش به نام‌های: بخش بابل‌کنار، بخش بندپی شرقی، بخش بندپی غربی، بخش گناب، بخش لاله آباد و بخش مرکزی می‌باشد.

به صورت افت شدید آب‌های زیرزمینی و کاهش دبی و حتی خشک شدن آب‌های سطحی در بسیاری از حوزه‌ها خودنمایی می‌کند (بوستانی، انصاری، ۱۳۸۹).

از آنجایی که از بین مصارف مختلف آب شهری (خانگی، تجاری، صنعتی، فضای سبز و غیره) بیش از ۶۱ درصد در کاربری مسکونی مورد استفاده قرار می‌گیرد، مطالعه مصرف آب در بخش مسکونی ضروری به نظر می‌رسد (مرکز آمار ایران، ۱۳۹۰). با توجه به حجم بالای اطلاعات مکانی و غیر مکانی در زمینه میزان مصرف آب و نیاز به استخراج الگوهای مصرف از بانک‌های اطلاعاتی، همچنین توانایی بالای روش‌های داده کاوی در تحلیل حجم زیادی از داده‌ها (Fayyad, Piatetsky-Shapiro, & Smyth, 1996; Han, Pei, & Kamber, 2011)

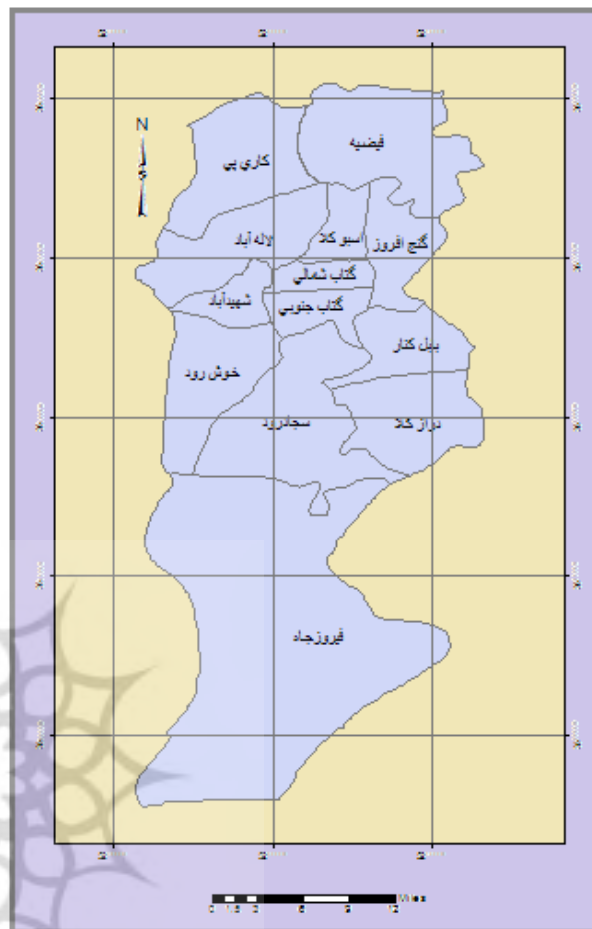
در این مطالعه به بررسی عوامل مرتبط با میزان مصرف آب شهری با استخراج قوانین وابستگی و استفاده از الگوریتم درخت تصمیم پرداخته شده است. بر این اساس، اهداف کلی مطالعه حاضر، بررسی و تحلیل الگوی مکانی مصرف آب خانگی در شهرستان بابل و همچنین بررسی و شناسایی عوامل مؤثر بر الگوی مصرف آب به وسیله استخراج قوانین وابستگی مبتنی بر برخی از پارامترهای مکانی و اجتماعی- اقتصادی از قبیل تعداد افراد خانواده، تعداد واحد، ویلایی بودن خانه‌ها، مساحت حیاط خانه و فضای سبز، فاصله از رودخانه بابلرود، فاصله از جاده‌های اصلی، تراکم جمعیت، درصد جوان بودن جمعیت و میزان مصارف آب در این شهرستان، تعیین نواحی کم‌مصرف و پر مصرف، بررسی توزیع مکانی مصرف آب در شهرستان بابل و در نهایت شناسایی و طبقه‌بندی عوامل مرتبط با مصرف آب خانگی در سطح محله‌های شهرستان بابل از طریق مدل درخت تصمیم می‌باشد.

بر این اساس مقاله حاضر در ۶ بخش به قرار ذیل تدوین شده است. در بخش اول به ارائه مقدمه‌ای در رابطه با کلیات تحقیق، اهمیت و میزان مصرف آب و همچنین مطالعات صورت گرفته در رابطه با مصرف آب و مدل‌سازی با استفاده از روش‌های داده‌کاوی پرداخته شد. در بخش‌های

فصلنامه علمی - پژوهشی اطلاعات جغرافیایی (ص ۵۷)
تحلیل مصرف آب شهری (خانگی) شهرستان بابل ... / ۵۷

حدود ۷۷/۶۳ درصد را دربر گرفته است (مرکز آمار ایران، ۱۳۹۰). بنابراین بررسی مصرف آب در این شهرستان و یافتن راه کارهای مناسب جهت کاهش مصرف آب در شهرستان بابل از اهمیت فراوان برخوردار است.

بین سال‌های ۱۳۹۵-۱۳۹۰، میانگین مصرف سالانه مشترکین خانگی در مناطق شهرستان بابل برابر با ۱۸۱/۲ متر مکعب در سال است و بخش‌های بندپی غربی، گتاب و بابل کنار بالاتر از میانگین و بخش‌های لاله آباد، مرکزی بابل و بندپی شرقی پایین‌تر از میانگین مصرف آب خانگی را به خود اختصاص داده‌اند. بخش لاله آباد با میانگین ۱۷۸/۸ متر مکعب، بخش مرکزی بابل با میانگین ۱۷۹/۲۸ متر مکعب و بخش بندپی شرقی با میانگین ۱۸۰/۲۴ متر مکعب پایین‌تر از حد میانگین و بخش‌های گتاب و بابل کنار با تفاوت کمی از حد میانگین به ترتیب دارای میانگین مصرف ۱۸۲/۱۶ و ۱۸۱/۶۸ متر مکعب می‌باشند. بخش بندپی غربی با اختلاف بیشتر از حد میانگین برابر با ۱۸۵/۰۴ متر مکعب در سال پرمصرف‌ترین بخش شهرستان بابل می‌باشد.



نگاره ۱: شهرستان بابل به تفکیک محله‌ها

۳- داده‌ها و روند کلی تحقیق

داده‌های مورد استفاده در این تحقیق شامل پارامترهای مکانی و اجتماعی-اقتصادی و اطلاعات میزان مصرف آب است. پارامترهای مکانی مؤثر در میزان مصرف آب مانند فاصله منازل از راه‌های اصلی، فاصله از رودخانه بابلرود، مساحت باغچه و فضای سبز در ساختمان و میزان عرصه و اعیان در ساختمان (مساحت حیاط خانه) در نظر گرفته شده‌اند. همچنین پارامترهای اجتماعی-اقتصادی از جمله تراکم جمعیت، ویلایی یا آپارتمانی بودن ساختمان‌ها، تعداد واحدهای مسکونی، متوسط تعداد افراد هر خانواده، و درصد پیر یا جوان بودن با توجه به داده‌های در دسترس از مرکز آمار ایران انتخاب شدند.

جهت محاسبه متوسط تعداد افراد خانواده در سطح محله، جمعیت هر محله بر تعداد خانواده آن تقسیم شده است. به منظور بررسی ویلایی و آپارتمانی بودن

یکی از عوامل رشد و توسعه اقتصادی هر منطقه‌ای را می‌توان وجود شرایط طبیعی مناسب از جمله منابع آب جهت رونق بخش کشاورزی و غیره نام برد (پورزیری *et al.*, ۱۳۹۵). رودخانه بابلرود با طول ۷۸ کیلومتر، تأمین کننده عمده منابع آب‌های سطحی و زیرزمینی شهرستان بابل به شمار می‌آید که آب آبیاری و شرب شهرستان را تأمین می‌کند. رودخانه بابلرود از رودهای پرآب استان در شهرستان بابل است که از کوه‌های البرز سرچشمه می‌گیرد و شاخه مهم آن سجادرود است. این رود پس از گذر از غرب شهرستان بابل و آبیاری شالیزارهای بابل، در بابلسر به دریای خزر می‌ریزد. آمار موجود بین سال‌های مختلف بیانگر این است که در شهرستان بابل مصرف آب خانگی

جدول ۱: دسته‌بندی میزان مصرف سالانه آب خانوار

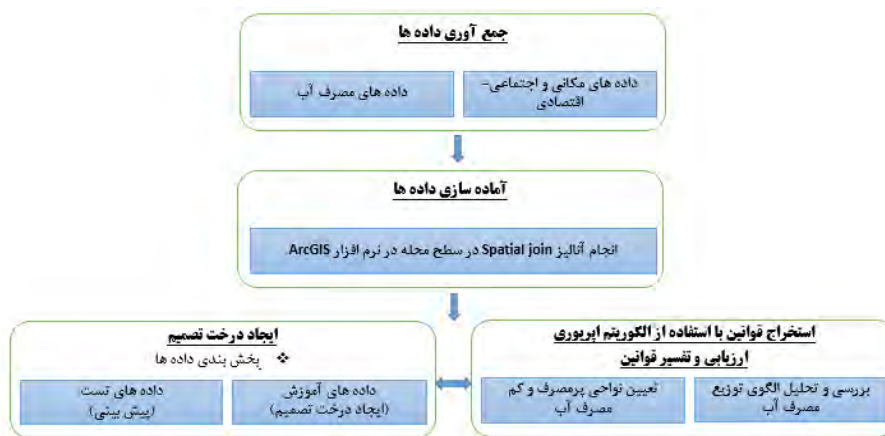
میزان مصرف (متر مکعب در سال)	دسته
۶۸۷-۸۰۴	مصرف خیلی کم
۸۰۴-۹۱۲	مصرف کم
۹۱۲-۹۸۸	مصرف متوسط
۹۸۸-۱۰۵۱	مصرف زیاد
۱۰۵۱-۱۲۵۴	مصرف خیلی زیاد

کاوش قوانین وابستگی دارای دو مرحله است. در مرحله اول، از هر پارامتر مؤثر در میزان مصرف آب برای ایجاد یک تراکنش استفاده می‌شود و در مرحله دوم، این تراکنش به تراکنش برداری تبدیل می‌گردد و تعدادی از قوانین وابستگی ارائه می‌شوند. در این مطالعه از نرم افزار KNIME به دلایلی همچون امکان ارتباط با پایگاه داده‌های مختلف، کاربرپسند بودن، ایجاد خروجی به فرمت دلخواه، بصری‌سازی مناسب برای داده‌ها و تحلیل‌ها و خروجی‌های موجود، به منظور استخراج قوانین وابستگی استفاده شده است. در ابتدا مقادیر مربوط به هر یک از پارامترها و میزان مصرف آب در این نرم افزار به عنوان ورودی و با فرمت اکسل فراخوانی شده است. در این مطالعه الگوریتم مورد استفاده برای استخراج قوانین، الگوریتم Apriori بوده است. تأثیر هر یک از پارامترهای مختلف به صورت جداگانه مورد بررسی قرار گرفت و هر یک از قوانین استخراج شده بیان کننده رابطه یکی از این پارامترها با میزان مصرف آب است. در میان قوانین به دست آمده، قوانین با support برابر ۲۵ و مقدار confidence برابر ۵۰ انتخاب شدند. در مرحله بعد، قوانین به دست آمده مورد ارزیابی قرار گرفته‌اند.

ارزیابی قوانین به دست آمده در ۲ بخش صورت گرفته است. در بخش اول، به تعیین نواحی پر مصرف و کم مصرف آب و بررسی الگوی مکانی توزیع مصرف آب خانوار پرداخته شده است. در بخش دوم، جهت ارزیابی قوانین به بررسی سهم هر یک از پارامترها در میزان مصرف آب خانوار پرداخته شده است. سپس به منظور طبقه‌بندی پارامترهای

ساختمان‌ها، مشترکینی با یک واحد، ویلایی و بیشتر از یک واحد، آپارتمانی در نظر گرفته شده‌اند. بعد از آماده سازی داده‌ها، مجموع میزان هر یک از پارامترها در سطح محله‌های شهرستان بابل با استفاده از آنالیز Spatial Join در نرم افزار ArcGIS محاسبه شده است. به منظور محاسبه مساحت حیاط خانه، میزان مساحت پارسل‌ها به عنوان عرصه در نظر گرفته شده است. پس از آنالیز Spatial Join، میزان عرصه و اعیان هر محله از هم کم شده است تا میزان مساحت حیاط خانه در سطح محله به دست آید. جهت به دست آوردن متغیر فاصله از جاده‌های اصلی، در هر یک از جاده‌های اصلی، در نرم افزار ArcGIS بافرهایی به شعاع ۵۰ متر، ۱۰۰ متر و ۲۰۰ متر تعریف شده است و مشترکینی که به ترتیب به فاصله هر یک از این شعاع‌ها هستند، مشخص شده‌اند.

داده‌های مربوط به میزان مصرف آب به صورت مجموع کل مصرف ۵ سال بین سال‌های ۱۳۹۰-۱۳۹۵ و در سطح هر مشترک در اختیار بوده است. به منظور محاسبه متوسط میزان مصرف، این مقادیر بر ۵ تقسیم شده‌اند. سپس با استفاده از آنالیز Spatial Join، مجموع میزان مصارف آب یک سال خانوار مانند سایر پارامترها در هر محله محاسبه شد. متوسط مصرف هر خانوار (میزان مصارف آب بر اساس واحدهای مصرف کننده) با تقسیم میزان مصرف هر محله بر تعداد واحدهای موجود به دست آمده است. در این مطالعه با استفاده از استخراج قوانین وابستگی، ارتباط میان پارامترهای مکانی و اجتماعی-اقتصادی و میزان مصرف آب کشف و بررسی شده است. به منظور استخراج قوانین وابستگی، این داده‌ها بر اساس مقدار در ۵ دسته طبقه‌بندی شده‌اند. کمترین میزان مصرف آب و مقدار هر یک از پارامترها در دسته‌های خیلی کم و کم قرار گرفته و به همین ترتیب با افزایش مقدار آنها در دسته‌های متوسط و زیاد و خیلی زیاد طبقه‌بندی شده‌اند. در جدول (۱) دسته‌بندی ایجاد شده برای میزان مصرف سالانه آب خانوار نشان داده شده است.



نگاره ۲: روند کلی تحقیق

بررسی عوامل مؤثر در میزان مصرف آب با توجه به توانایی بالای این روش‌ها در حل مسائل پیچیده و تحلیل حجم بالای داده‌ها، می‌تواند بسیار کمک کننده باشد.

۴-۱- داده کاوی^۱

افزایش سریع توانایی‌های فنی بشر برای تولید و جمع‌آوری داده‌ها و همچنین استفاده همگانی از وب و اینترنت به عنوان یک سیستم اطلاع‌رسانی جهانی، بشر را با حجم زیادی از داده‌ها و اطلاعات مواجه می‌کند. این رشد سریع در داده‌های ذخیره شده، ضرورت استفاده از تکنولوژی‌های جدید و ابزارهای خودکار که به صورت هوشمند به انسان کمک کنند تا این داده‌ها را به اطلاعات و دانش تبدیل کند بیش از پیش ضروری می‌نماید. داده‌کاوی یا اکتشاف دانش از پایگاه داده به معنای فرآیند استخراج اطلاعاتی است که قبلاً بر ما پوشیده بوده و قابلیت استفاده و ارزش زیادی دارند. کشف دانش در پایگاه‌داده فرآیند شناسایی درست، ساده و مفید الگوها و مدل‌های قابل فهم در داده‌ها می‌باشد. به عبارت دیگر، داده‌کاوی به فرآیند استخراج دانش مفید از داده اطلاق می‌شود (نگاره ۳). داده‌ها اغلب حجیم اما بدون ارزش می‌باشند و داده به تنهایی قابل استفاده نیست، بلکه دانش نهفته در داده‌ها قابل استفاده می‌باشد (Fayyad et al., 1996; Han et al., 2011).

مؤثر در مصرف و همچنین پیش‌بینی میزان مصرف آب بخشی از مشترکین یک درخت تصمیم ایجاد شد. برای ساخت درخت تصمیم، داده‌ها به دو دسته داده‌های آموزش و داده‌های تست تقسیم شدند. حدود ۷۰ درصد از کل داده‌ها به عنوان داده‌های آموزش و ۳۰ درصد باقیمانده به عنوان داده‌های آزمون به مدل معرفی شدند. در مدل درخت تصمیم، پارامتر وابسته (هدف)، میزان مصرف آب و پارامترهای مکانی و اجتماعی-اقتصادی به عنوان پارامترهای مستقل در نظر گرفته شدند. بعد از تقسیم داده‌ها به عنوان داده‌های آموزش و تست، یک درخت تصمیم طبقه‌بندی کننده ایجاد گردید. سپس مدل درخت تصمیم به مدل مجموعه قوانین، تبدیل و طبقه‌بندی انجام شده توسط درخت تصمیم به صورت قوانین مستقل بیان شده است. در مرحله آخر، میزان مصرف آب با استفاده از درخت تصمیم ایجاد شده پیش‌بینی گردید و دقت پیش‌بینی متغیر هدف برای داده‌های تست محاسبه شد. نگاره (۲) روند کلی تحقیق را نشان می‌دهد.

۴- مبانی نظری تحقیق

به دلیل اهمیت یافتن عوامل مؤثر در مصرف آب و چگونگی توزیع مکانی مصرف شهروندان، امروزه پژوهشگران به استفاده از روش‌های نوین برای مدل‌سازی میزان مصرف و عوامل مؤثر بر آن ترغیب شده‌اند. استفاده از روش‌های داده‌کاوی و تحلیل‌های مکانی مختلف برای

در یک تراکنش خرید زیاد است، دارای وابستگی هستند و کشف مجموعه عناصر تکرار شونده به کشف وابستگی بین این عناصر در مجموعه داده‌ها منجر می‌شود. در این فرآیند، با توجه به اقلام مختلفی که مشتریان در سبد خریدشان قرار می‌دهند، رفتار خرید آن‌ها مورد تحلیل قرار می‌گیرد (Agrawal et al., 1993; Ali, Manganaris, & Srikanth, 1997). بررسی ارزش و معیار مقبولیت قوانین وابستگی، دو متغیر پشتیبان^۲ و اطمینان^۳ قوانین مورد استفاده قرار می‌گیرند. پشتیبان درصدی از مجموعه تراکنش‌های پایگاه داده (N) است که شامل X و Y می‌باشند. از طرفی، اطمینان، درصدی از مجموعه تراکنش‌هایی است که اگر شامل X هستند شامل Y نیز باشند (X و Y هر کدام یک مجموعه از اقلام می‌باشند، همچنین هر تراکنش شامل مجموعه‌ای از اقلام است) (رابطه ۱) (Agarwal



نگاره ۳: داده کاوی به عنوان یک مرحله از فرآیند کشف

دانش (Fayyad et al., 1996; Han et al., 2011)

(Srikant, 1994; Agrawal et al., 1993; Ali et al., 1997)

$$Support = \frac{freq(X,Y)}{N} \quad (1)$$

$$Confidence = \frac{freq(X,Y)}{freq(X)}$$

Dataset

Transaction ID	Items
1	Shoes, Shirt, Jacket
2	Shoes, Jacket
3	Shoes, Jeans
4	Shirt, Sweatshirt

Frequent Itemsets

Frequent Itemset	Support
{Shoes}	75%
{Shirt}	50%
{Jacket}	50%
{Shoes, Jacket}	50%

Rules

Frequent Itemset	Support	Confidence
Shoes ⇒ Jacket	50%	66.6%
Jacket ⇒ Shoes	50%	100%

نگاره ۴: نمونه‌ای از قوانین وابستگی (Ali et al., 1997)

در نگاره (۴)، نمونه‌ای از قوانین وابستگی نشان داده شده است. این پایگاه داده شامل ۴ تراکنش است و جدول سوم شامل قوانین استخراج شده است.

در سال‌های اخیر داده‌کاوی به دلایل حجم بالای داده‌ها، سهولت دسترسی به کامپیوتر و سهولت به کارگیری، موارد استفاده زیادی یافته است (Fayyad et al., 1996; Han et al., 2011). داده‌کاوی، روش‌ها و الگوریتم‌های فراوانی دارد. در ادامه دو مورد از روش‌های داده‌کاوی همچون کاوش قوانین وابستگی و درخت تصمیم بررسی خواهند شد.

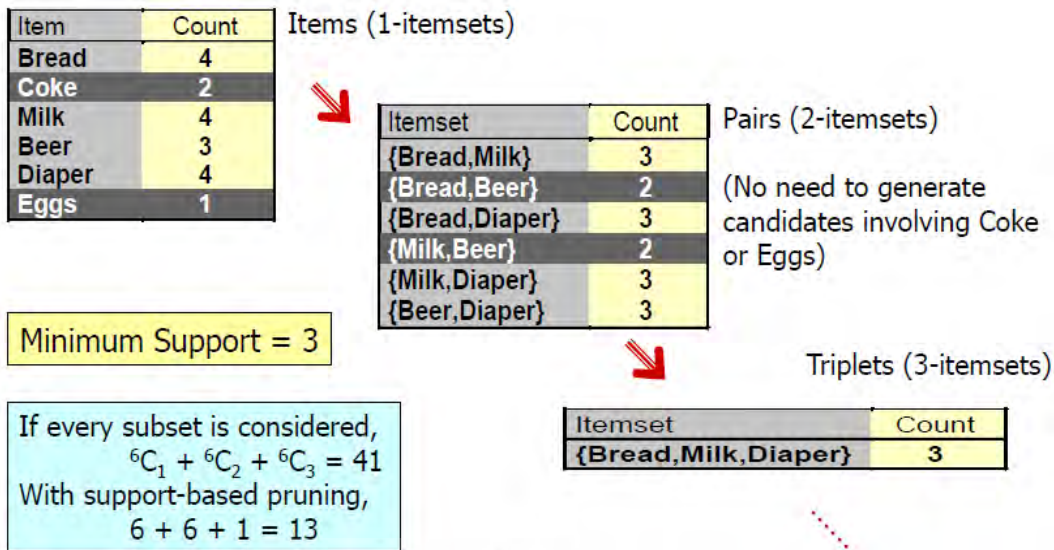
۴-۲- کاوش قوانین وابستگی^۱

از این روش می‌توان به عنوان یک روش در مبحث داده‌کاوی به منظور کشف ارتباطات مهم و با اهمیت در پایگاه‌داده‌های بزرگ بهره برد. کشف روابط وابستگی بین حجم عظیم تراکنش‌های کسب و کار می‌تواند در تشخیص تقلب، در حوزه پزشکی، در روش‌های به کارگیری وب توسط کاربران و شخصی‌سازی و از همه مهمتر، در طراحی و پیاده‌سازی سیستم‌های توصیه گر مورد استفاده قرار گیرد (M.-S. Chen, Park, & Yu, 1996). قوانین وابستگی، روابط متقابل بین مجموعه بزرگی از اقلام داده‌ای را نشان می‌دهند. مثال متداول در خصوص کشف قوانین وابستگی، تحلیل سبد خرید است. اجناسی که در یک فروشگاه احتمال خرید آنها

2- Support
3- Confidence

1- Association Rules Mining

فصلنامه علمی - پژوهشی اطلاعات جغرافیایی (مهر)
تحلیل مصرف آب شهری (خانگی) شهرستان بابل ... / ۶۱



نگاره ۵: مراحل ایجاد مجموعه اقلام مهم با استفاده از الگوریتم Apriori (Agrawal et al., 1996; Tan, 2006)

است (Dastourani et al., 2013). این روش در زمینه‌های مختلفی از جمله طبقه‌بندی گیاهان، تشخیص بیماری و استراتژی‌های بازاریابی به کار می‌رود. این ساختار تصمیم‌گیری با استفاده از روش‌های ریاضیاتی و محاسباتی به توصیف و طبقه‌بندی یک مجموعه از داده‌ها کمک می‌کند.

از طرفی، ابزاری مفید برای تجزیه و تحلیل‌های مقدماتی و تولید فرضیات در مناطقی که با کمبود اطلاعات است، به شمار می‌آید. با استفاده از درخت‌های تصمیم می‌توان خصوصیات کمی و کیفی را به طور همزمان پردازش نمود. بنابراین درخت تصمیم نشان‌دهنده یک مدل منطقی از پدیده مورد تحقیق می‌باشد (Geissen, Kampichler, López-de Llergo- Juárez, & Galindo-Acántara, 2007). یک درخت معمولاً از ریشه، شاخه‌ها، گره‌ها (جایی که شاخه‌ها منشعب می‌شوند) و برگ‌ها تشکیل شده است. درخت تصمیم را معمولاً از چپ به راست یا از بالا به پایین ترسیم می‌کنند به طوری که ریشه در بالا قرار بگیرد. گره اول را ریشه می‌گویند. از هر یک از گره‌های داخلی (غیر برگ)، به دو یا چند شاخه دیگر می‌توانند منشعب شوند و هر گره نمایشگر یک خصوصیت معین می‌باشد و هر گره داخلی نیز با یک ویژگی مشخص می‌شود. برگ‌های درخت با یک کلاس و یا یک دسته از

۴-۲-۱- کاوش قوانین وابستگی با استفاده از الگوریتم Apriori

معروف‌ترین الگوریتم در کاوش قوانین وابستگی، الگوریتم Apriori است. برتری اصلی این الگوریتم در مقایسه با سایر الگوریتم‌ها، حجم محاسباتی کمتر آن است (Tan, 2006). این الگوریتم از اصل Apriori تبعیت می‌کند که براساس آن اگر یک مجموعه آیتم پرتکرار باشد همه زیرمجموعه‌های آن نیز پرتکرار خواهند بود. به عبارت دیگر، بر اساس این اصل، پشتیبان هر مجموعه آیتم هرگز بزرگتر از پشتیبان زیرمجموعه‌اش نیست (رابطه ۲). با استفاده از این اصل می‌توان مجموعه آیتم‌های پرتکرار را هرس نموده و قوانین وابستگی قوی‌تری را استخراج کرد. (Agrawal, Mannila, Srikant, Toivonen, & Verkamo, 1996; Tan, 2006)

نگاره (۵) مثالی از ایجاد مجموعه اقلام مهم با استفاده از الگوریتم Apriori را نشان می‌دهد.

$$\forall X, Y : (X \subseteq Y) \Rightarrow s(X) \geq s(Y) \quad (2)$$

۴-۳- درخت تصمیم

درخت تصمیم یکی از روش‌های داده‌کاوی به شمار می‌آید که ابزار قدرتمند و رایجی برای طبقه‌بندی و پیش‌بینی

لازم است که ابتدا خودهمبستگی مکانی بین موارد وقوع رویداد بررسی شود و تعیین گردد که الگوی مکانی رویداد مورد نظر در منطقه از کدام توزیع (تصادفی، یکنواخت یا خوشه‌ای) پیروی می‌کند. در مورد خودهمبستگی مکانی فرض صفر بیان می‌کند که هیچ الگوی مکانی میان موارد وقوع رویداد مورد نظر وجود ندارد.

فرض صفر، یک نوع فرض مورد استفاده در آمار است و بیانگر آن است که هیچ اهمیت آماری در یک مجموعه از مشاهدات داده وجود ندارد. فرض صفر تا زمانی پذیرفته است که شواهد آماری، آن را برای یک فرض جایگزین باطل نکنند (AhangarCani et al., 2016; Narayanan, 1993). از این رو به منظور بررسی و تشخیص وجود خودهمبستگی مکانی، آزمون‌های مختلفی از جمله شاخص Global Moran's I پیشنهاد شده‌اند (AhangarCani et al., 2016; Getis, Anselin, Lea, Ferguson, & Müller, 2004; Mitchell, 1999; Ruiz, Tedesco, McTighe, Austin, & Kitron, 2004)

۴-۴-۱- شاخص عمومی خودهمبستگی Global Moran's I

با استفاده از شاخص عمومی Global Moran's I می‌توان درجه پراکنده بودن یا متمرکز بودن داده‌های مکانی را در فضا اندازه‌گیری نمود. مقادیر شاخص موران بین -۱ تا +۱ متغیر است. به طور کلی، مقادیر مثبت، بیانگر خودهمبستگی مثبت و مقادیر منفی بیانگر خودهمبستگی منفی می‌باشند. چنانچه مقدار شاخص موران برابر صفر باشد نشان‌دهنده عدم وجود خودهمبستگی مکانی می‌باشد. اگر مقدار شاخص به عدد +۱ نزدیک باشد، داده‌ها دارای الگوی خوشه‌ای و اگر مقدار شاخص به عدد -۱ نزدیک باشد، داده‌ها از هم گسسته و پراکنده هستند. مقدار شاخص عمومی Global Moran's I را می‌توان با استفاده از رابطه (۳) محاسبه نمود (AhangarCani et al., 2016; Mitchell, 2005; Moran, 1950)

$$I = \frac{N}{\sum_{i=1}^N \sum_{j=1}^N w_{ij}} * \frac{\sum_{i=1}^N \sum_{j=1}^N W_{ij} (x_i - \bar{X})(x_j - \bar{X})}{\sum_{i=1}^N (x_i - \bar{X})^2} \quad (3)$$

جواب‌ها مشخص می‌شوند و هر برگ درخت، یک کلاس یا یک دسته را مشخص می‌کند (Breiman, Friedman, Olshen, & Stone, 1984; Geissen et al., 2007)

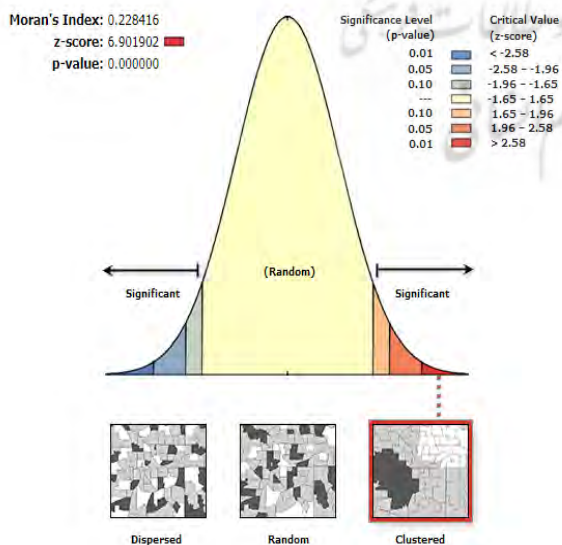
۴-۴-۲ آمار فضایی

تحلیل آماری کلاسیک نقش اساسی در تکنیک‌های داده‌کاوی ایفا می‌کند. به طوری که آمار، اساس اغلب تکنولوژی‌هایی می‌باشد که داده‌کاوی روی آن‌ها بنا می‌شود. در مطالعات محیطی اغلب با داده‌ها و مشاهداتی سر و کار داریم که مستقل از یکدیگر نیستند و نوعاً وابستگی آنها ناشی از مکان قرار گرفتن مشاهدات در فضای مورد مطالعه می‌باشد. از این رو، در مطالعه اینگونه مشاهدات نمی‌توان از روش‌های سنتی آمار بهره گرفت، چرا که این داده‌ها دارای ساختاری پیوسته در مکان و زمان می‌باشند.

در مطالعات علوم محیطی، این داده‌ها را داده‌های فضایی می‌نامند. از جمله ویژگی‌های این داده‌ها این است که مشاهدات نزدیک وابستگی بیشتری با هم دارند و برعکس. در واقع می‌توان بیان نمود که وابستگی مشاهدات تابع فاصله موقعیت مشاهدات می‌باشد. لذا مطالعات داده‌های فضایی نیازمند روشی به‌هنگار جهت پاسخ رفتار این داده‌ها در مکان و زمان می‌باشند (AhangarCani, Farnaghi, & Shirzadi, 2016; Jolliffe & Philipp, 2010; Tobler, 1970)

بر مبنای چنین نیازی، روش‌های آماری معمول برای واکاوی چنین داده‌هایی مناسب نیستند. زیرا فرض اصلی بیشتر روش‌های معمول آماری بر اساس استقلال داده‌ها می‌باشد. از این رو، به دلیل وجود همبستگی و اثر مکانی میان این قسم از داده‌ها، این فرض در حقیقت محقق نشده و داده‌ها به هم وابسته هستند. بنابراین لازم است به نحوی ساختار همبستگی داده‌ها در تحلیل آن‌ها لحاظ گردد. بدین منظور آمار فضایی به عنوان گزینه‌ای مناسب برای تحلیل این قسم از داده‌ها می‌تواند مورد بررسی قرار گیرد (AhangarCani et al., 2016; ENTEZARI, DADASHI, & ASADI, 2016; Kendall, 1998; Ripley, 2005). بنابراین به منظور مدل‌سازی برخی از رویدادها

دارای مصرف آب بیشتری نسبت به سایر مناطق شهرستان هستند و جزء مناطق پرمصرف می‌باشند و با رنگ زرد مشخص شده‌اند. در این محله‌ها، ساختمان‌های مسکونی عموماً ویلایی و دارای مساحت حیاط و فضای سبز بالاتر و تراکم جمعیت پایین تر نسبت به سایر محله‌ها می‌باشند. این عوامل به نوعی می‌توانند نشان‌دهنده بالاتر بودن سطح درآمد خانواده‌ها نیز باشد. همچنین، مناطق جنوبی شهرستان دارای کمترین میزان مصرف آب می‌باشند. دلیل چنین نتیجه‌ای را این‌طور می‌توان تفسیر نمود که ساکنان مناطق جنوبی شهرستان بابل، غالباً از اهالی روستا می‌باشند و از سطح درآمد پایین‌تری برخوردارند و همچنین برای مصارف غیر شرب نیز از آب‌های چاه استفاده می‌کنند. نتایج بررسی وجود خودهمبستگی در الگوی توزیع مصرف آب نشان می‌دهند که شاخص موران برابر ۰/۲۲۸۴ و z-score برابر ۶/۹۰۱۹ و p-value برابر صفر شده است. همانطور که در نگاره (۷) مشخص است، این داده‌ها دارای خودهمبستگی مکانی و دارای الگوی خوشه‌ای هستند. بنابراین خوشه‌ای بودن الگوی توزیع مصرف آب بیانگر دخالت داشتن برخی از عواملی می‌باشند که باعث می‌شوند میزان مصرف آب در مناطقی از شهرستان بابل بیشتر از سایر مناطق باشد.

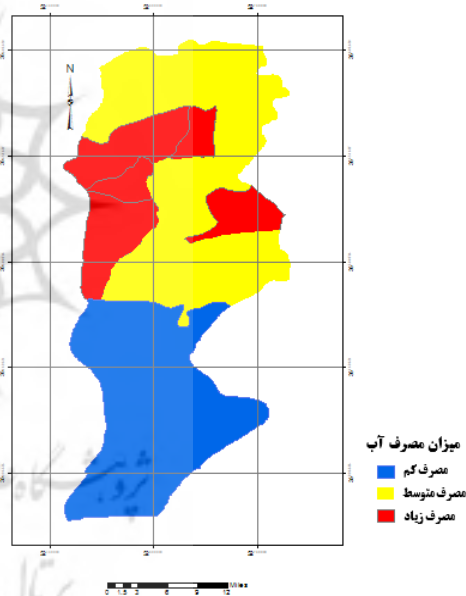


نگاره ۷: بررسی الگوی توزیع مصرف آب خانوار شهرستان بابل با استفاده از شاخص Moran's I

که در آن x_j و x_i به ترتیب تعداد موارد وقوع رویداد (میزان مصرف آب) در i امین و j امین واحد مکانی (محله) می‌باشند. \bar{x} بیانگر متوسط تعداد موارد وقوع رویداد در هر محله می‌باشد. N تعداد محله‌ها در شهرستان بابل را نشان می‌دهد و W_{ij} نشان‌دهنده وزن مکانی میان موارد i و j می‌باشد.

۵- نتایج و بحث

در این قسمت از تحقیق، نتایج حاصل از بررسی عوامل مرتبط با مصرف آب خانوار مورد بررسی قرار می‌گیرد. این بررسی در سه مرحله تحلیل مکانی میزان مصرف و بررسی قوانین وابستگی و درخت تصمیم انجام می‌شود.



نگاره ۶: توزیع مکانی مصرف آب سالانه خانوارها در شهرستان بابل

۵-۱- مصرف آب خانگی شهرستان بابل

نگاره (۶)، میزان مصرف آب شهرستان بابل را برای هر محله نشان می‌دهد. همانطور که در این نگاره نمایش داده شده است منازل مسکونی در محله‌های غربی و در ساحل رودخانه بابلرود، مصرف آب بالاتری داشته و با رنگ قرمز مشخص شده‌اند. علاوه بر حاشیه رودخانه، محله‌هایی که در مرکز، شمال و شمال شرقی شهرستان قرار دارند

جدول ۲- قوانین استخراج شده با استفاده از الگوریتم Apriori

پارامتر مورد بررسی	درجه اهمیت در درخت تصمیم	نوع رابطه با مصرف خانوار	سهم هر پارامتر در غیاب سایر پارامترها (R ²)	قوانین استخراج شده
ویلايي بودن ساختمان	-	مستقیم	-	۱- آپارتمانی \rightarrow مصرف خانوار خیلی کم ۲- مصرف خانوار کم \rightarrow ویلايي ۳- مصرف خانوار متوسط \rightarrow ویلايي
مساحت حیاط خانه	اولین درجه اهمیت	مستقیم	۸۵ درصد	۴- مصرف خانوار خیلی زیاد \rightarrow ابعاد حیاط خیلی زیاد ۵- مصرف خانوار زیاد \rightarrow ابعاد حیاط زیاد ۶- مصرف خانوار کم \rightarrow ابعاد حیاط خیلی کم ۷- مصرف خانوار خیلی کم \rightarrow ابعاد حیاط خیلی کم
درصد جوان بودن جمعیت	سومین درجه اهمیت	معکوس	۲۱ درصد	۸- مصرف خانوار زیاد \rightarrow درصد جوان کم ۹- مصرف خانوار خیلی زیاد \rightarrow درصد جوان متوسط ۱۰- درصد جوان زیاد \rightarrow مصرف خانوار کم ۱۱- درصد جوان خیلی زیاد \rightarrow مصرف خانوار کم
مساحت فضای سبز	-	مستقیم	۹۷ درصد	۱۲- مصرف خانوار خیلی کم \rightarrow ابعاد خیلی کم ۱۳- ابعاد فضای سبز متوسط \rightarrow مصرف متوسط ۱۴- مصرف خانوار زیاد \rightarrow ابعاد فضای سبز زیاد
تراکم جمعیت	-	معکوس	۸۲ درصد	۱۵- مصرف خانوار زیاد \rightarrow تراکم خیلی کم ۱۶- مصرف خیلی زیاد \rightarrow تراکم خیلی کم ۱۷- تراکم متوسط \rightarrow مصرف کم ۱۸- مصرف خیلی کم \rightarrow تراکم زیاد
تعداد واحد	دومین درجه اهمیت	معکوس	۶۴ درصد	۱۹- مصرف خانوار خیلی کم \rightarrow تعداد واحد خیلی کم ۲۰- تعداد واحد زیاد \rightarrow مصرف کم ۲۱- تعداد واحد کم \rightarrow مصرف کم ۲۲- تعداد واحد متوسط \rightarrow مصرف متوسط
تعداد افراد خانواده	دومین درجه اهمیت	مستقیم	۱۰۰ درصد	۲۳- مصرف خانوار خیلی کم \rightarrow تعداد افراد خیلی کم ۲۴- تعداد افراد زیاد \rightarrow مصرف خانوار کم
فاصله تا رودخانه بابلرود	اولین درجه اهمیت	معکوس	۹۲ درصد	۲۵- فاصله تا رودخانه خیلی زیاد \rightarrow مصرف کم ۲۶- فاصله تا رودخانه زیاد \rightarrow مصرف کم ۲۷- فاصله تا رودخانه خیلی کم \rightarrow مصرف زیاد ۲۸- فاصله تا رودخانه کم \rightarrow مصرف متوسط ۲۹- مصرف خانوار متوسط \rightarrow فاصله تا رودخانه متوسط
فاصله از جاده اصلی	-	مستقیم	۹۵ درصد	۳۰- فاصله از جاده ۱۰۰ متر \rightarrow مصرف خانوار کم ۳۱- مصرف خیلی کم \rightarrow فاصله از جاده ۵۰ متر ۳۲- مصرف متوسط \rightarrow فاصله از جاده ۲۰۰ متر

۲-۵- قوانین وابستگی به دست آمده با الگوریتم Apriori

افزایش پیدا می‌کند. همچنین با افزایش فاصله از رودخانه بابلرود، بافت منطقه تغییر می‌کند. در نتیجه میزان مصرف آب خانوار کاهش یافته و محله‌هایی با کمترین فاصله از رودخانه که عموماً محله‌های ساحلی هستند دارای بالاترین میزان مصرف می‌باشند. قوانین استخراج شده نیز مشخص کننده این واقعیت هستند که یک رابطه مستقیم بین دو متغیر مساحت حیاط خانه با میزان مصرف وجود دارد. این روابط می‌توانند مؤید مصرف آب برای نظافت حیاط خانه باشند. مطابق با قوانین ۱۲ تا ۱۴ در جدول (۲)، این

این بخش از تحقیق، قوانین کشف شده با استفاده از الگوریتم Apriori را ارائه می‌دهد. قوانین مرتبط با هر ویژگی مکانی و اجتماعی - اقتصادی به تفکیک پارامتر مورد بررسی در جدول (۲) ارائه شده‌اند.

همانطور که در جدول (۲) مشخص شده است، با توجه به قوانین استخراج شده از الگوریتم Apriori، با افزایش فاصله از راه‌های اصلی، میزان مصرف آب خانوار

میزان مصرف آب خانوار افزایش می‌یابد. در این راستا، اگر خانواده‌ها به مصرف بهینه در زمینه شستشو و نظافت ترغیب شوند می‌توان در آب مصرفی خانوار صرفه‌جویی نمود. قوانین ۱۸-۱۵ نشان می‌دهند که میزان مصرف آب با تراکم جمعیت رابطه عکس دارد و با بالا رفتن تراکم جمعیت، میزان مصرف آب خانوار کاهش می‌یابد.

۳-۵- درخت تصمیم

عواملی که در درخت تصمیم در شاخه‌های بالاتر قرار می‌گیرند دارای اهمیت بیشتری می‌باشند و تغییر آنها تأثیر بیشتری در میزان مصرف دارد. هر چه به سمت پایین درخت تصمیم حرکت می‌کنیم اهمیت این عوامل کاهش می‌یابد. از میان عوامل معرفی شده به درخت تصمیم، به ترتیب، مساحت حیاط خانه و فاصله از رودخانه بابلرود، اهمیت بیشتری نسبت به سایر عوامل دارند و سپس تعداد واحد و تعداد افراد خانواده از نظر اهمیت در رده بعد قرار می‌گیرند (نگاره ۸). همانطور که در جداول (۲) و (۳) قابل مشاهده است، با مقایسه قوانین به دست آمده از درخت تصمیم و قوانین به دست آمده از الگوریتم Apriori مشاهده می‌شود که این دو مجموعه قوانین در بسیاری از موارد با هم همپوشانی دارند. به عنوان مثال، در هر دو مجموعه قانون به

قوانین مؤید این مطلب است که میزان مساحت باغچه و فضای سبز خانگی با مصرف آب رابطه مستقیم دارد. از طرفی مطابق با جدول، میزان مصرف آب با درصد جوان بودن جمعیت رابطه عکس دارد و با افزایش درصد جوان جمعیت، میزان مصرف آب کاهش می‌یابد. علت این رابطه را می‌توان تبلیغات انجام شده در سال‌های اخیر در مورد نحوه صحیح مصرف دانست. از آنجایی که جمعیت نسل جوان با شیوه‌های صحیح مصرف آب بیشتر آشنا می‌باشند، از اینرو، از هدر رفت آب بهتر جلوگیری می‌کنند. قوانین ۱۹، ۲۱ و ۲۲ بیانگر این حقیقت هستند که با افزایش تعداد واحد، میزان مصرف آب سالانه هر خانوار در سطح محله افزایش می‌یابد. نکته مهم در این قوانین مربوط به قانون ۲۰ است که نشان می‌دهد که اگر تعداد واحدها از تعداد مشخصی بالاتر باشد میزان مصرف به ازای هر واحد کاهش می‌یابد.

این کاهش می‌تواند ناشی از مشترک بودن حیاط، فضای سبز، پارکینگ و راه پله و در واقع کاهش سرانه این فضاها باشد و در نتیجه میزان آب مصرفی مربوط به آنها بین واحدها سرشکن می‌شود و میزان مصرف آب هر خانوار کاهش می‌یابد. در رابطه با تعداد افراد خانواده نیز می‌توان بیان نمود که یک رابطه مستقیم بین مصرف آب و تعداد افراد خانواده وجود دارد و با افزایش تعداد افراد خانواده



نگاره ۸: ساختار کلی درخت تصمیم ایجاد شده جهت طبقه بندی عوامل مرتبط با میزان مصرف خانوار

۶- نتیجه گیری و ارائه پیشنهادات

این مطالعه با استفاده از الگوریتم Apriori به استخراج قوانین وابستگی جهت کشف ارتباط میان میزان مصرف آب و برخی از پارامترهای مکانی و اجتماعی-اقتصادی مانند فاصله از راه‌های اصلی، فاصله از رودخانه بابلرود، تعداد واحدهای مسکونی، مساحت فضای سبز و باغچه خانگی، تراکم جمعیت، متوسط تعداد افراد هر خانواده، جوان بودن جمعیت، ویلایی یا آپارتمانی بودن خانه‌ها و مساحت حیاط در خانه پرداخته است. همچنین با ایجاد درخت تصمیم، عوامل مرتبط با مصرف طبقه‌بندی شده و با استفاده از آن میزان مصرف آب مشترکین پیش‌بینی شده است.

قوانین به دست آمده با دقت ۷۵ درصد، میزان مصرف آب داده‌های آزمون را پیش‌بینی نموده‌اند. این مدل امکان تحلیل و آینده‌نگری در میزان مصرف آب و در نتیجه برنامه‌ریزی در جهت کاهش مصرف و مدیریت منابع آب را فراهم می‌سازد.

برخی از پارامترهای مورد بررسی مانند تعداد افراد خانوار، مساحت باغچه خانگی، مساحت حیاط خانه و فاصله از جاده‌های اصلی با میزان مصرف آب رابطه مستقیم و برخی دیگر مانند فاصله از رودخانه بابلرود، تراکم جمعیت، آپارتمانی بودن ساختمان‌ها، تعداد واحدهای مسکونی و جوان بودن جمعیت با میزان مصرف آب رابطه معکوس دارند.

به منظور ارزیابی قوانین به دست آمده، به بررسی تعیین نواحی پرمصرف و کم‌مصرف و همچنین بررسی خودهمبستگی در الگوی توزیع مصرف آب پرداخته شده است. مناطق جنوبی شهرستان به عنوان نواحی با کمترین میزان مصرف آب مشخص شدند. محله‌های غربی و ساحلی، مرکزی، شمالی و شمال شرقی، به ترتیب دارای بیشترین میزان مصرف هستند.

در این محله‌ها، خانه‌ها ویلایی و دارای تراکم جمعیت پایین‌تر و مساحت فضای سبز و حیاط بالاتری نسبت به سایر محله‌ها هستند که بالا بودن میزان مصرف در آنها قابل

دست آمده از درخت تصمیم و الگوریتم Apriori، اگر تعداد افراد خانواده در دسته خیلی زیاد و فاصله از رودخانه در دسته زیاد قرار داشته باشد، میزان مصرف کم خواهد بود. در درخت تصمیم با ثابت بودن درصد جوان جمعیت و تعداد افراد خانواده، اگر تعداد واحد از متوسط به زیاد تغییر کند، میزان مصرف آب خانوار نیز از کم به متوسط افزایش می‌یابد. این روند در قوانین به دست آمده از الگوریتم Apriori نیز قابل مشاهده است و با تغییر واحد از کم به متوسط میزان مصرف خانوار از کم به متوسط افزایش می‌یابد و در مورد پارامتر مساحت حیاط خانه در هر دو مجموعه قانون با افزایش مساحت حیاط خانه میزان مصرف افزایش می‌یابد. قوانین به دست آمده با دقت ۷۵ درصد، میزان مصرف آب داده‌های آزمون را پیش‌بینی نموده‌اند.

جدول ۳: قوانین به دست آمده از درخت تصمیم ایجاد شده

جهت پیش‌بینی میزان مصرف آب
درصد جوان جمعیت متوسط و تعداد افراد زیاد \rightarrow مصرف آب کم
درصد جوان جمعیت زیاد و تعداد افراد زیاد \rightarrow مصرف آب کم
درصد جوان جمعیت متوسط و تعداد افراد خیلی زیاد و فاصله از رودخانه زیاد \rightarrow مصرف آب خیلی کم
درصد جوان جمعیت خیلی زیاد و تعداد افراد زیاد \rightarrow مصرف آب خیلی کم
درصد جوان جمعیت خیلی کم و تعداد افراد زیاد و تعداد واحد متوسط \rightarrow مصرف آب کم
درصد جوان جمعیت خیلی کم و تعداد افراد زیاد و تعداد واحد زیاد \rightarrow مصرف آب متوسط
درصد جوان جمعیت کم و تعداد افراد خیلی زیاد و مساحت حیاط کم \rightarrow مصرف آب کم
درصد جوان جمعیت کم و تعداد افراد زیاد و مساحت حیاط خیلی کم و تراکم متوسط \rightarrow مصرف آب خیلی کم
فضای سبز متوسط و درصد جوان جمعیت زیاد و فاصله از رودخانه متوسط \rightarrow مصرف آب متوسط
فضای سبز کم و درصد جوان جمعیت زیاد و فاصله از رودخانه خیلی زیاد \rightarrow مصرف آب خیلی کم
فضای سبز خیلی زیاد و درصد جوان جمعیت متوسط و فاصله از رودخانه متوسط \rightarrow مصرف آب متوسط
فضای سبز زیاد و درصد جوان جمعیت متوسط و فاصله از رودخانه زیاد \rightarrow مصرف آب کم

منابع و مأخذ

- ۱- آفاحسینعلی شیرازی، اکبرپور؛ محسن، ابوالفضل، (۱۳۹۰)، برآورد تقاضای روزانه آب شهری با استفاده از سری فوریه: مطالعه موردی شهر بیرجند در استان خراسان جنوبی، کنفرانس بین المللی آب و فاضلاب، تهران - شرکت مهندسی آب و فاضلاب کشور.
- ۲- ابراهیمی، نادری؛ پرویز، حسن، (۱۳۸۰)، بررسی و ارزیابی مدیریت عرضه و تقاضای آب شرب در شرایط خشکسالی اصفهان، مجله آب و محیط زیست (۴۹-۴۸)، ۸۹-۹۷.
- ۳- بوستانی، انصاری؛ آرمین، حسین، (۱۳۸۹)، ارائه رویکردهای مناسب در مدیریت مصرف آب شهری، همایش علمی چالش آب در استان قم گذشته، حال و آینده، دانشگاه قم.
- ۴- جوادیانزاده، محمدمهدی، (۱۳۸۸)، تهیه تابع تقاضای آب شهری با استفاده از روش شبکه های عصبی مصنوعی در شهر یزد، سومین همایش ملی آب و فاضلاب (با رویکرد اصلاح الگوی مصرف) ۴ تا ۵ اسفند ۱۳۸۸، تهران - پژوهشگاه نیرو.
- ۵- شمسایی، م. (۱۳۷۹). برآورد تابع تقاضای آب استان اصفهان، مجموعه ۲۱ مقاله ارائه شده در اولین همایش علمی و تحقیقاتی بهینه سازی مصرف آب، انتشارات روابط عمومی و امور بین الملل شرکت آب و فاضلاب استان تهران.
- ۶- شهرستانی، حسین، (۱۳۹۳)، سازماندهی و مدیریت مصرف بهینه آب در بخش کشاورزی. فصلنامه نظام مهندسی کشاورزی و منابع طبیعی، ۴۵(۱۲)، ۳۷-۴۱.
- ۷- مرکز آمار ایران، (۱۳۹۰)، سالنامه آماری کشور (<https://salnameh.sci.org.ir/AllUser/DirectoryTreeComplete.aspx>).
- ۸- محمدی، ر. (۱۳۹۳). بررسی تأثیر هدفمندی یارانه‌ها بر الگوی مصرف و میزان مصرف آب شهر اردبیل. پایان نامه دوره کارشناسی ارشد، دانشکده مدیریت، دانشگاه آزاد اسلامی واحد گرمی، اردبیل.

انتظار است. نتایج بررسی وجود خودهمبستگی در الگوی توزیع مصرف آب نشان می‌دهند که شاخص موران برابر با ۰/۲۲۸۴ و z-score برابر با ۶/۹۰۱۹ و p-value برابر صفر شده است و داده‌ها دارای خودهمبستگی مکانی و دارای الگوی خوشه‌ای هستند.

در این تحقیق، با مقایسه قوانین به دست آمده از الگوریتم Apriori و قوانین به دست آمده از درخت تصمیم مشخص می‌شود که این دو مجموعه قانون در بسیاری از موارد با هم همپوشانی دارند و از میان عوامل ذکر شده به ترتیب، مساحت حیاط خانه، فاصله از رودخانه بابلرود، تعداد واحد و تعداد افراد خانواده در طبقه‌بندی درخت تصمیم در شاخه‌های بالاتر قرار دارند و در نتیجه دارای اهمیت بیشتری نسبت به سایر عوامل می‌باشند.

به عنوان کارهای آینده، از آنجایی که در این تحقیق به علت عدم دسترسی به اطلاعات، برخی از پارامترها همچون هزینه مصرف آب و وضعیت تحصیلی افراد خانواده در نظر گرفته نشده‌اند، پیشنهاد می‌شود که پارامترهای اجتماعی-اقتصادی بیشتری مورد بررسی قرار گیرند و به کشف پارامترهای بیشتر در مصرف آب شهری پرداخته شود. همچنین مصرف آب خانوار به صورت دوره ۵ ساله در دسترس بوده است که در مطالعات آینده با اطلاعات و داده‌های کامل‌تر و با در نظر گرفتن یک دوره زمانی بیشتر می‌توان نتایج بهتر و جامع‌تری به دست آورد. در این مطالعه به دلیل محدودیت زمانی و داده‌ای، مدلسازی به صورت سالانه انجام گرفته است. لذا در مطالعات آینده با توجه به اینکه میزان مصرف آب در ماه‌ها و فصول مختلف متفاوت می‌باشد و متغیر زمان در میزان آب مصرفی خانوار تأثیر می‌گذارد، پیشنهاد می‌شود که مدل‌سازی زمان‌مند به صورت ماهانه و یا فصلی انجام شود تا از این طریق بتوان به مدیریت بهینه مصرف آب کمک نمود.

- 19- Chen, Park & Yu, M.-S, J.-S, P. S. (1996). Data mining for path traversal patterns in a web environment. *Distributed Computing Systems, Proceedings of the 16th International Conference on 1996 May 27, IEEE*, pp. 385-392.
- 20- Dastourani, Habibipoor, Ekhtesasi, Talebi & Mahjoobi, M., A., M., A., J. (2013). Evaluation of the Decision Tree model in Precipitation Prediction (case study: Yazd synoptic station). *Iran Water Resources Research*, 8(3), 14-27.
- 21- Entezari, Dadashi & Asadi, A., R.A., M. (2016). Assessing the Autocorrelation of Spatial - Temperature change in heat islands of Khorasan Razavi Province, pp. 39-43.
- 22- Fallahi, Ansari & Moghaddas, M., H., S. (2012). Evaluating effective factors on household water consumption and forecasting its demand: panel data approach. *Water Waste J*, 23(4), 78-87.
- 23- Fayyad, Piatetsky-Shapiro & Smyth, U., G., P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, 17(3), 37.
- 24- Geissen, Kampichler, López-de Llergo-Juárez & Galindo-Acántara, V., C., J., A. (2007). Superficial and subterranean soil erosion in Tabasco, tropical Mexico: development of a decision tree modeling approach. *Geoderma*, 139(3-4), 277-287.
- 25- Getis, Anselin, Lea, Ferguson & Miller, A., L., A., M., H. (2004). *Spatial analysis and modeling in a GIS environment. A research agenda for geographic information science*. CRC, Boca Raton, FL, 157-196.
- 26- Han, Pei & Kamber, J., J., M. (2011). *Data mining: concepts and techniques*: Elsevier.
- 27- Jolliffe & Philipp, I. T., A. (2010). Some recent developments in cluster analysis. *Physics and Chemistry of the Earth, Parts A/B/C*, 35(9-12), 309-315.
- 28- Kendall, W. S. (1998). *Perfect simulation for the area-interaction point process*, Probability towards 2000, Springer, New York, pp. 218-234.
- 29- Kheir, Chorowicz, Abdallah & Dhont, R. B., J., C., D. (2008). Soil and bedrock distribution estimated from gully form and frequency: a GIS-based decision-tree model for Lebanon. *Geomorphology*, 93(3-4), 482-492.
- ۹- ملکی نسب، ا.، ابریشم‌چی، ا.، و تجریشی، م. (۱۳۸۶). ارزیابی صرفه‌جویی در مصرف آب خانگی به واسطه استفاده از قطعات کاهنده مصرف. *مجله آب و فاضلاب*، ۱۸(۲)، ۱۱-۲.
- ۱۰- ورزیری، وکیل پور، مرتضوی؛ آزیتا، محمدحسن، سید ابوالقاسم، (۱۳۹۵). بررسی اثر قیمت‌گذاری اقتصادی آب آبیاری بر الگوی کشت در دشت دهگلان. *فصلنامه علمی - پژوهشی تحقیقات اقتصاد کشاورزی*، ۸(۳۱)، ۸۱-۱۰۰.
- 11- Agarwal & Srikant, R., R. (1994). Fast algorithms for mining association rules. in *Proc. 20th int. conf. very large data bases, VLDB*, pp. 487-499.
- 12- Agrawal, Imieliński & Swami, R., T., A. (1993). Mining association rules between sets of items in large databases. *Acm sigmod record (ACM)*, Vol. 22, No. 2, pp. 207-216.
- 13- Agrawal, Mannila, Srikant, Toivonen & Verkamo, R., H., R., H., A. I. (1996). Fast discovery of association rules. *Advances in knowledge discovery and data mining*, 12(1), 307-328.
- 14- AhangarCani, Farnaghi, & Shirzadi, M., M., M. (2016). Predictive Map of Spatio-Temporal Distribution of Leptospirosis Using Geographical Weighted Regression and Multilayer Perceptron Neural Network Methods. *Journal of Geomatics Science and Technology*, 6(2), 79-98.
- 15- Ali, Manganaris & Srikant, K., S., R. (1997). Partial Classification Using Association Rules. In *KDD*, vol. 97, pp. p115-118.
- 16- Appice, Ceci, Lanza, Lisi & Malerba, A., M., A., F. A., D. (2003). Discovery of spatial association rules in geo-referenced census data: A relational mining approach. *Intelligent Data Analysis*, 7(6), 541-566.
- 17- Breiman, Friedman, Olshen & Stone, L., J., R., C. (1984). *Classification and Regression Trees (CART)* Wadsworth. Pacific Grove, CA.
- 18- Chen, Shu, Ning & Chen, J.-C, C.-S, S.-K, H.-W. (2008). Flooding probability of urban area estimated by decision tree and artificial neural networks. *Journal of Hydroinformatics*, 10(1), 57-67.

(2012). Seasonal and annual regional drought prediction by using data-mining approach. *Atmósfera*, 25(1), 85-105.

30- Kim & Yum, Y. S., B.-J , (2011), Recommender system based on click stream data using association rule mining. *Expert Systems with Applications*, 38(10), 13320-13327.

31- Mitchell, A. (1999). *The ESRI guide to GIS analysis: geographic patterns & relationships*, Volume 1, ESRI, Inc.

32- Mitchell, A. (2005). *The ESRI guide to GIS analysis*, Volume 2, spatial measurements and statistics ESRI Press. Redlands CA: Esri Press.

33- Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37(1/2), 17-23.

34- Narayanan, A. (1993). *Probability and Statistics in Engineering and Management Science*: Taylor & Francis, pp. 238-239.

35- Ripley, B. D. (2005). *Spatial statistics (Vol. 575)*: John Wiley & Sons.

36- Ruiz, Tedesco, McTighe, Austin & Kitron, M. O., C., T. J., C., U. (2004). Environmental and social determinants of human risk during a West Nile virus outbreak in the greater Chicago area, 2002. *International Journal of Health Geographics*, 3(1), 8.

37- Sabouhi & NOUBAKHT, M., M. (2009). Estimating the water demand function of Pardis city. *Journal of Water and Wastewater*, 20, pp. 69-74.

38- Statistical Center of Iran, (2011), *Statistical Pocketbook of the Islamic Republic of Iran*, Management & Planning Organization, Tehran. 786.

39- Taghi Sattari, Anli, Apaydin & Kodal, M., A., H., S. (2012). Decision trees to determine the possible drought periods in Ankara. *Atmósfera*, 25(1), 65-83.

40- Tan, P.-N. (2006). *Introduction to data mining*: Pearson Education India.

41- Tobler, W. R., (1970), A computer movie simulating urban growth in the Detroit region. *Economic geography*, 46(sup1), 234-240.

42- Versichele, De Groote, Bouuaert, Neutens, Moerman & Van de Weghe, M., L., M. C. T., I., N. (2014). Pattern mining in tourist attraction visits through association rule learning on Bluetooth tracking data: A case study of Ghent, Belgium. *Tourism Management*, 44, 67-81.

43- Yurekli, Taghi Sattari, Anli & Hinis, K., M., A., M.



پروہشگاہ علوم انسانی و مطالعات فرہنگی
پرتال جامع علوم انسانی