

پیش‌بینی وقوع بارش روزانه با استفاده از داده‌های هواشناسی روزهای قبل (مطالعه موردی: شهر اصفهان)

قربان مهتابی^۱

استادیار گروه مهندسی آب، دانشکده کشاورزی دانشگاه زنجان، زنجان، ایران

فرشید تاران

دکتری آبیاری و زهکشی، گروه مهندسی آب، دانشکده کشاورزی دانشگاه تبریز، تبریز، ایران

سعید مظفری

دانشجوی کارشناسی ارشد منابع آب و عضو بنیاد ملی نخبگان، گروه مهندسی آب، پردیس ابوریحان، دانشگاه تهران، تهران، ایران

تاریخ دریافت: ۱۳۹۶/۶/۲۶ تاریخ پذیرش: ۱۳۹۷/۲/۲۹

چکیده

هدف از این پژوهش، پیش‌بینی وقوع بارش روزانه شهر اصفهان با استفاده از داده‌های هواشناسی ۱ تا ۷ روز قبل است. برای این منظور، داده‌های هواشناسی دوره ۲۰۰۹-۲۰۰۰ با استفاده از مدل‌های هوشمند بردار پشتیبان، k- نزدیک‌ترین همسایگی، شبکه عصبی مصنوعی و درخت تصمیم بررسی گردید. نتایج نشان داد که در هر چهار روش، دقت پیش‌بینی بهترین سناریوها با استفاده از داده‌های ۶ و ۷ روز قبل، کمتر از ۷۵ درصد بود، اما با استفاده از داده‌های روزهای ۱ تا ۵ روز قبل، بارش روزانه با دقت بیش از ۸۰ درصد پیش‌بینی شد. عملکرد روش درخت تصمیم بهتر از سه روش دیگر بود و به علت ارائه درخت تصمیم‌گیری، نتایج سناریوهای ۱ تا ۵ روز قبل این روش ارائه شد. نتایج سناریوها با استفاده از داده‌های ۱ تا ۳ روز قبل نشان داد که رطوبت نسبی هوا مناسب‌ترین پارامتر برای پیش‌بینی وقوع بارش روزانه است، اما در شرایط استفاده از داده‌های ۴ و ۵ روز قبل، دمای هوا مناسب‌ترین پارامتر برای انجام پیش‌بینی بود. در نهایت عملکرد بهترین سناریوها با استفاده از داده‌های دوره ۲۰۱۶-۲۰۱۰ صحت‌سنجی شد. بهترین نتایج در بخش صحت‌سنجی به ترتیب مربوط به سناریوی ۱ روز قبل (با پارامتر حداقل رطوبت نسبی) و سناریوی ۴ روز قبل (با پارامتر دمای حداکثر) بود.

واژگان کلیدی: بارش روزانه، درخت تصمیم، شبکه عصبی، بردار پشتیبان، k-نزدیک‌ترین همسایگی.

مقدمه

بارش از مهم‌ترین پارامترهای اقلیمی در حوزه مدیریت منابع آب و از اصلی‌ترین داده‌های ورودی به سیستم‌های

هیدرولوژیکی به شمار می‌آید. پیش‌بینی این پدیده برای اهداف مختلفی از قبیل تخمین سیلاب، خشک‌سالی، آبخیزداری، گردشگری، کشاورزی و برنامه‌ریزی آبیاری اهمیت فراوانی دارد (فریوان و سیگیز اوغلو^۱، ۲۰۰۵). با اطلاع از زمان وقوع و مقدار بارش می‌توان امکان وقوع سیل را در منطقه بررسی و تمهیدات و پیشگیری‌های لازم را به عمل آورد. همچنین با اطلاع به موقع از زمان وقوع بارش در طول سال می‌توان برنامه‌ریزی مناسبی برای انجام بهینه فعالیت‌های کشاورزی و گردشگری انجام داد. با وجود پیشرفت‌های تکنولوژی، اگرچه پیش‌بینی وقوع بارش در دامنه زمان و مکان امری محال نیست، اما پیچیدگی‌های خاص خود را دارد، چرا که تابع عوامل بسیاری از جمله دما و رطوبت است. مدل‌های مفهومی و آماری مختلفی برای پیش‌بینی و پیش‌آگاهی متغیرهای اقلیمی ارائه شده است، اما امروزه سیستم‌های هوشمند مبتنی بر کاوش داده‌ها در مدل‌سازی فرایندهای هیدرولوژیکی و مهندسی آب مورد توجه قرار گرفته‌اند؛ به عبارت دیگر، این سیستم‌ها به عنوان روش‌هایی معتبر و شناخته شده در مدل‌سازی فرایندهای پیچیده غیرخطی قادرند با استفاده از داده‌های هواشناسی و هیدرولوژیکی، تخمین دقیق‌تری از شرایط منطقه ارائه دهند (بهاتاچاریا و سولوماتین^۲، ۲۰۰۵). در سال‌های اخیر، انواع سیستم‌های هوشمند در بخش‌های مختلف مهندسی آب، به ویژه در زمینه پیش‌بینی پارامترهای هواشناسی و هیدرولوژیکی مورد استفاده قرار گرفته‌اند.

پارتال و سیگیز اوغلو^۳ (۲۰۰۹) به منظور پیش‌بینی بارش روزانه شهرهای مختلف ترکیه، از روش شبکه عصبی-موجک بهره گرفتند. نتایج نشان داد که این روش ترکیبی، مقادیر بارش‌های صفر و حداکثر را با دقت بالایی برآورد می‌کند. ویراسینگ^۴ و همکاران (۲۰۱۰) عملکرد کمک شبکه عصبی در پیش‌بینی مقدار و وقوع بارش روزانه بررسی کردند. نتایج نشان داد که دقت پیش‌بینی روزهای بدون بارش بهتر از روزهای با بارش بود. وو^۵ و همکاران (۲۰۱۰) برای پیش‌بینی بارش ماهانه در ایستگاه‌های مختلف دنیا، از چهار روش شبکه عصبی مدولار^۶ (MANN)، شبکه عصبی مصنوعی (ANN)، روش نزدیک‌ترین همسایگی^۷ (KNN) و رگرسیون خطی^۸ استفاده و گزارش نمودند که با وجود عملکرد مناسب مدل MANN، کارایی مدل شبکه عصبی مصنوعی بهتر از مدل MANN است. آنیهوتری و موهاپاترا^۹ (۲۰۱۲) به مطالعه پیش‌بینی وقوع باران‌های موسمی روزانه تابستان در منطقه کارناتاکای هند پرداختند. آن‌ها با استفاده از روش رگرسیون پیش‌رونده و پارامترهای هواشناسی روزانه، گزارش کردند که مدل احتمالاتی پیشنهادی دارای دقت ۵۷ الی ۹۰ درصد در پیش‌بینی شرایط بارش روزانه است. مکانیک^{۱۰} و همکاران (۲۰۱۳) با بررسی پیش‌بینی بارش فصل بهار در حوضه ویکتوریای استرالیا با استفاده از روش‌های شبکه عصبی و رگرسیون چندگانه، نشان دادند که روش شبکه

1- Freiwan & Cigizoglu

2- Bhattacharya & Solomatine

3- Partal & Cigizoglu

4- Weerasinghe

5- Wu

6- Modoular Neroul Network

7- K-Nearest Neighbourhood

8- Linear Regression

9- Agnihotri & Mohapatra

10- Mekanik

عصبی مصنوعی، عملکرد بهتری در پیش‌بینی بارش دارد. ماندال و چودهاری^۱ (۲۰۱۴) با استفاده از سری مدل‌های احتمالاتی، احتمال وقوع مقادیر بارش‌های روزانه حداکثر را پیش‌بینی کردند. بر اساس نتایج، در جزیره ساگار هند، احتمال وقوع مجموع بارش ماهانه بیش از ۱۰۰ میلی‌متر در فصل‌های تابستان و زمستان در حدود ۷۸-۸۵ درصد به‌دست آمد. جورج^۲ و همکاران (۲۰۱۶) با استفاده از یک مدل دوبعدی خطی، مقدار بارش روزانه را بر اساس داده‌های هواشناسی همان روز پیش‌بینی کردند. نتایج نشان داد که مدل می‌تواند با دقت ۸۰ درصد، مقادیر بارش را پیش‌بینی نماید.

فقیه (۱۳۸۹) تلفیق شبکه عصبی مصنوعی^۳ و الگوریتم ژنتیک^۴ در پیش‌بینی بارش ماهانه، نتیجه گرفتند که روش تلفیقی، عملکرد بهتری نسبت به روش شبکه عصبی مصنوعی دارد. فلاحی و همکاران (۱۳۹۰) عملکرد مدل درخت تصمیم^۵ (DT) را در پیش‌بینی مقدار بارش سه ایستگاه هواشناسی بندرانزلی، اراک و قم مطالعه کردند که نتایج، بیانگر توانایی و دقت بالای این مدل در پیش‌بینی مقدار بارش بود. ستاری و همکاران (۱۳۹۳) از دو روش شبکه عصبی و مدل درخت تصمیم M5 به‌منظور پیش‌بینی بارش ماهانه ایستگاه اهر استفاده نمودند. نتایج نشان داد که هر دو روش توانایی خوبی در پیش‌بینی بارش ماهانه دارند. بابا علی و دهقانی (۱۳۹۶) کارایی سه روش برنامه‌ریزی بیان ژن^۶، سیستم استنتاج فازی-عصبی^۷ و شبکه عصبی مصنوعی را در برآورد بارش ماهانه حوضه کاکا رضا بررسی کردند و اظهار داشتند که مدل برنامه‌ریزی بیان ژن، توانایی بالایی در تخمین بارش ماهانه دارد. بهرامی و همکاران (۱۳۹۶) با بررسی تأثیر روش‌های پیش‌پردازش^۸ در عملکرد روش شبکه عصبی مصنوعی، نتیجه گرفتند که روش پیش‌پردازش مینیمم-ماکزیمم، بهترین عملکرد را در پیش‌بینی بارش ماهانه ایستگاه آباده دارد. همچنین شبکه با دو پارامتر ورودی حداقل دما و رطوبت نسبی نتایج خوبی نسبت به سایر تعداد ورودی‌های مختلف داشت.

همان‌طور که اشاره شد، پیش‌بینی وقوع بارش به‌ویژه در مقیاس روزانه، با استفاده از داده‌های هواشناسی روز یا روزهای قبل، کمک شایانی به مدیریت سیل، برنامه‌ریزی انواع فعالیت‌های کشاورزی، گردشگری و غیره می‌کند. مرور پژوهش‌های پیشین نشان می‌دهد که مطالعات خاصی در زمینه پیش‌بینی وقوع بارش روزانه بر مبنای داده‌های هواشناسی با کمک انواع مدل‌های هوشمند صورت نگرفته است. هدف این پژوهش، مطالعه پیش‌بینی وقوع بارش روزانه به کمک مدل‌های هوشمند بردار پشتیبان، k-نزدیک‌ترین همسایگی، شبکه عصبی مصنوعی و درخت تصمیم است. برای این منظور، از داده‌های روزانه هواشناسی شهر اصفهان استفاده شد و وقوع یا عدم وقوع بارش روزانه با استفاده از داده‌های هواشناسی روزهای قبل بررسی و بهترین مدل پیش‌بینی شرایط بارش روزانه ارائه گردید.

1- Mandal & Choudhury

2- George

3- Artificial Neural Network

4- Genetic Algorithm

5- Decition Tree

6- Gene expression programing

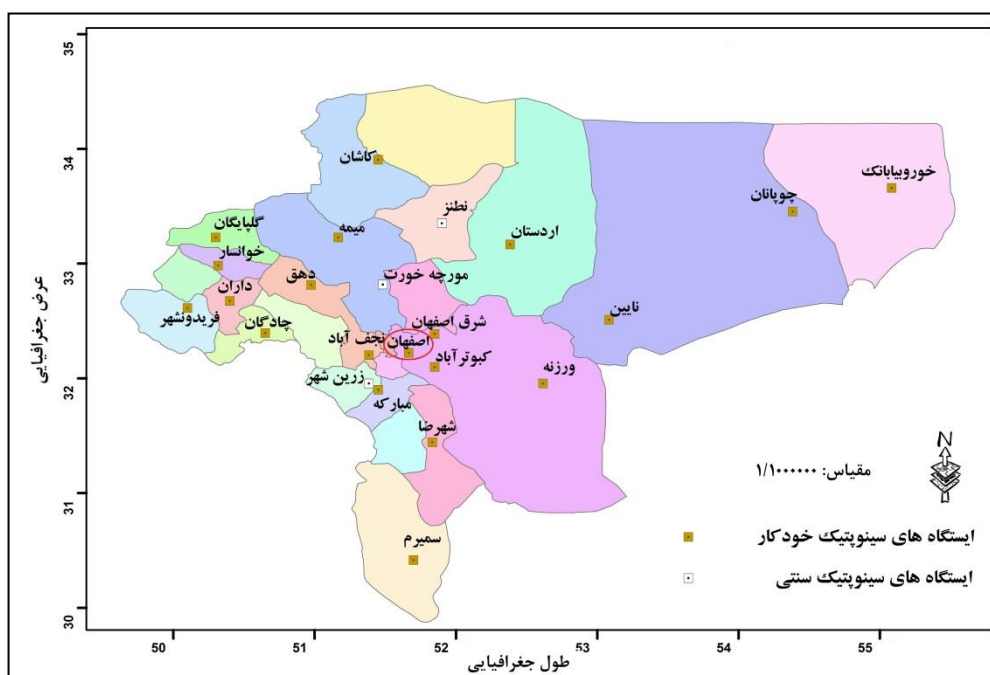
7- Fuzzy-Neural Inference System

8- Preprosseing

داده‌ها و روش‌ها

منطقه مورد مطالعه و داده‌های مورد استفاده

در این پژوهش، از داده‌های هواشناسی ایستگاه سینوپتیک اصفهان استفاده شد. کلان‌شهر اصفهان به‌عنوان سومین شهر بزرگ ایران و مرکز استان اصفهان، از مهم‌ترین مراکز گردشگری و صنعتی کشور بوده و دارای اقلیم گرم و خشک است. موقعیت جغرافیایی ایستگاه سینوپتیک هواشناسی اصفهان در طول $40^{\circ} 51'$ و عرض $37^{\circ} 32'$ و ارتفاع 1550 متر از سطح دریا قرار دارد (شکل ۱).



منبع: سایت اداره کل هواشناسی استان اصفهان

شکل ۱: موقعیت ایستگاه هواشناسی شهر اصفهان

در این پژوهش، از داده‌های روزانه هواشناسی ۱۰ سال آماری مربوط به سال‌های ۲۰۰۹-۲۰۰۰ استفاده شده است. داده‌های روزانه شامل مجموع بارش (R)، دمای متوسط هوا (T_{mean})، دمای حداقل (T_{min})، دمای حداکثر (T_{max})، متوسط رطوبت نسبی (RH_{mean})، حداقل رطوبت نسبی (RH_{min})، حداکثر رطوبت نسبی (RH_{max})، متوسط سرعت باد ($Wind$)، درصد ابرناکی ($Cloud$) و مجموع ساعات آفتابی (Sun) هستند. بر مبنای داده‌های بارش، روزهای سال در دو کلاس روزهای وقوع بارش (Y) و روزهای عدم وقوع بارش (N) دسته‌بندی شدند و داده‌های هواشناسی در قالب ۷ سری داده با ۱ تا ۷ روز تأخیر برای پیش‌بینی شرایط وقوع یا عدم وقوع بارش روزانه بررسی گردید.

روش‌های مورد استفاده

۱- روش طبقه‌بندی کننده بردار پشتیبان (SVC)

روش بردار پشتیبان، یک روش یادگیری با نظارت است که برای طبقه‌بندی و پیش‌بینی به کار می‌رود. هدف روش طبقه‌بندی بردار پشتیبان، یافتن ابر صفحه‌ای با بهترین توانایی تعمیم برای تفکیک داده‌های مربوط به دوطبقه با حاشیه (فاصله بین ابر صفحه و نزدیک‌ترین داده آموزشی) حداکثر است (واپنیک، ۱۹۹۵)^۱. در این روش برای به دست آوردن نتایج دقیق، انتخاب تابع کرنل بهینه که خصوصیات غیرخطی را به صورت خطی درمی‌آورد، حائز اهمیت است. از مهم‌ترین توابع کرنل می‌توان به چندجمله‌ای ساده و نرمال شده، کرنل شعاع‌محور و پیرسون اشاره نمود. هر کدام از این توابع کرنل، شامل پارامترهایی به نام هایپرپارامتر هستند که علاوه بر اهمیت انتخاب تابع کرنل بهینه، تعیین مقدار هایپرپارامترهای آن نیز بسیار مهم است.

۲- روش k- نزدیک‌ترین همسایگی (KNN)

الگوریتم k- نزدیک‌ترین همسایگی، یک روش طبقه‌بندی بر اساس سنجش فاصله است. در این روش، ابتدا داده‌های آموزشی و طبقه‌های متناظر با آن‌ها در نظر گرفته می‌شود. سپس برای طبقه‌بندی یک نمونه جدید، فاصله آن با هر یک از نمونه‌های آموزشی محاسبه شده و K همسایه از میان نزدیک‌ترین همسایه‌ها (نمونه‌های آموزشی) انتخاب می‌گردند. در نهایت، نمونه جدید در طبقه‌ای قرار می‌گیرد که بیشترین نمونه‌های موجود در K همسایه‌اش در آن طبقه هستند. برای محاسبه فاصله هر نمونه جدید از نمونه‌های مشاهده‌ای، از توابع فاصله مانند همینگ، اقلیدسی و چبیشف استفاده می‌شود که در این مطالعه، تابع اقلیدسی مورد استفاده قرار گرفت (یانگ سو، ۲۰۱۱)^۲. در این الگوریتم برای دستیابی به بهترین نتیجه، تعیین مقدار بهینه پارامتر K اهمیت زیادی دارد که برای این منظور، از روش سعی و خطا استفاده شد.

۳- روش شبکه عصبی مصنوعی (ANN)

شبکه‌های عصبی مصنوعی، جزو سیستم‌های دینامیکی بوده و با تجزیه و تحلیل داده‌های ورودی و نتایج نظیر آن‌ها و بدون در نظر گرفتن فرآیند فیزیکی حاکم بر سیستم، قادر به ایجاد رابطه بین متغیر وابسته و مستقل هستند. به طور کلی، یک شبکه عصبی شامل سه لایه ورودی، پنهان و خروجی است. تعداد نرون‌های موجود در لایه‌های ورودی و خروجی، با توجه به نوع مسئله مشخص می‌شود. تعداد نرون‌های موجود در لایه‌های پنهان و تعداد این لایه‌ها از طریق سعی و خطا مشخص می‌گردد. کاربردی‌ترین نوع شبکه عصبی، شبکه‌های پرسپترون چندلایه (MLP) با قانون یادگیری پس انتشار پیش‌خور است که در این پژوهش نیز از این نوع شبکه استفاده شد. در این روش، داده‌ها پس از آماده‌سازی، استاندارد

1- Vapnik

2- Yang Su

می‌شوند. در طراحی شبکه، برای رسیدن به حالت بهینه، پارامترهای تعداد لایه‌های پنهان، نرخ یادگیری در لایه‌های پنهان و خروجی، مومنتوم و زمان یادگیری به روش سعی و خطا تغییر داده می‌شوند (مقدم نیا^۱ و همکاران، ۲۰۰۹).

۴- روش درخت تصمیم (DT)

درخت تصمیم، یک روش کارآمد برای طبقه‌بندی سلسله‌مراتبی است. برای ایجاد یک درخت تصمیم، سری داده به زیرمجموعه‌های همگن تقسیم می‌شوند، به طوری که ساختمان درختی آن، اطلاعاتی در رابطه با قابلیت داده‌ها برای پیش‌بینی یا قاعده‌سازی یک موضوع ارائه می‌دهد (پال و مثر، ۲۰۰۳)^۲. فرایند ایجاد یک درخت تصمیم شامل سه مرحله است: تقسیم گره‌ها، تعیین گره‌های انتهایی، تخصیص طبقه‌ها به گره‌های انتهایی. یک درخت تصمیم، ترکیبی از یک ریشه یا گره اصلی (دربرگیرنده تمام داده‌ها)، یک سری گره داخلی (شاخه‌ها) و یک سری گره‌های انتهایی (برگ‌ها) است. یکی از مهم‌ترین الگوریتم‌های درخت تصمیم، الگوریتم معروف و پرکاربرد C4.5 (روش J48) است که یک الگوریتم تک متغیره است (ویتن و فرانک، ۲۰۰۵)^۳. این الگوریتم، ناهمگنی داده‌ها را کاهش داده و موقعیت شاخه‌ها در هر گره داخلی درخت را به دقت تخمین می‌زند. روش J48 از معیاری به نام نسبت بهره اطلاعات^۴، برای تعیین آنتروپی (بی‌نظمی) دسته داده‌های تولید شده در اثر تقسیم‌بندی استفاده می‌کند. در این روش، برای کاهش خطای طبقه‌بندی داده‌های خارج از زیرمجموعه آموزشی، از یک روش هرس مبتنی بر خطا استفاده می‌شود (ویتن و فرانک، ۲۰۰۵).

در این پژوهش، برای آموزش و صحت‌سنجی روش‌های مورد بررسی، از روش اعتبارسنجی متقابل ۱۰ گانه^۵ استفاده شد (ویتن و فرانک، ۲۰۰۵). بدین منظور، داده‌ها به طور تصادفی به ۱۰ قسمت تقسیم می‌شوند، به طوری که هر طبقه تقریباً همان تناسب متناظر با کل داده‌ها را دارد. هر قسمت به صورت نوبتی کنار گذاشته شده (داده‌های آزمایشی) و آموزش بر روی ۹ قسمت دیگر انجام می‌گیرد. سپس میزان خطا برای داده‌های آزمایشی محاسبه می‌شود. فرایند آموزش ۱۰ مرتبه انجام شده و سرانجام از ۱۰ خطای به دست آمده، میانگین‌گیری می‌شود تا یک خطای کل به دست آید (ویتن و فرانک، ۲۰۰۵). در این پژوهش، به منظور ارزیابی دقت پیش‌بینی روش‌ها از آماره "درصد موارد پیش‌بینی صحیح"^۶ (CCI) استفاده گردید. آماره CCI متداول‌ترین معیار برای ارزیابی دقت مدل‌های طبقه‌بندی کننده است و هرچه مقدار آن به صد نزدیک باشد، دقت مدل بیشتر می‌گردد (ثاقبیان و همکاران، ۲۰۱۳).

یافته‌های پژوهش

به منظور پیش‌بینی وقوع بارش روزانه با استفاده از داده‌های هواشناسی روزهای قبل، سناریوهای مختلف پارامترهای

1- Moghaddamnia

2- Pal & Mather

3- Witten & Frank

4 Information gain ratio

5 10-fold cross-validation

6- Correctly Classified Instances

ورودی در روش‌های مورد نظر بررسی گردید. سناریوها در قالب ۷ سری داده با ۱ تا ۷ روز تأخیر در نظر گرفته شدند. در بررسی عملکرد انواع سناریوها با ترکیب پارامترهای مختلف (سناریوهای یک یا چند پارامتری)، با توجه به تعداد زیاد سناریوها، تنها نتایج سناریوهایی ارائه شد که درصد موارد پیش‌بینی صحیح (CCI) آن‌ها نزدیک به ۸۰ درصد و بیشتر بود. لازم به ذکر است که بررسی عملکرد سناریوها در حالت تأخیرهای ۶ و ۷ روزه نشان داد که درصد پیش‌بینی صحیح بهترین سناریوها برای هر چهار روش، کمتر از ۷۵ درصد بود و نتایج این دو حالت ارائه نشد. در جدول ۱، عملکرد سناریوهای مختلف در پیش‌بینی شرایط بارش ۱ روز بعد بر اساس سری داده‌های هواشناسی روزانه (تأخیر ۱ روزه پارامترها) برای روش‌های مختلف ارائه شده است. در هر سناریو، بهترین روش پیش‌بینی با علامت یک زیرخط و در هر روش، بهترین سناریو با علامت دو زیرخط مشخص شده است. مقایسه مقادیر درصد موارد پیش‌بینی صحیح (CCI) در هر روش نشان می‌دهد که تفاوت قابل توجهی بین عملکرد سناریوهای مختلف وجود ندارد. همچنین در هر روش، درصد پیش‌بینی صحیح سناریوهای ترکیبی و تک‌پارامتری اختلاف فاحشی باهم ندارند. در بین سناریوهای تک‌پارامتری، پارامترهای ساعات آفتابی، ابرناکی، حداکثر دمای روزانه، میانگین و حداقل رطوبت نسبی، به‌عنوان مهم‌ترین پارامترها برای پیش‌بینی شرایط بارش روز بعد به دست آمدند. با توجه به تأثیر غالب پارامترهای هواشناسی مذکور در پیش‌بینی شرایط بارش روز بعد، ترکیب پارامترها افزایش قابل توجهی در عملکرد سناریوهای ترکیبی نداشت. شایان ذکر است که نتایج سناریوهای ترکیبی با بیش از دو پارامتر، به علت دارا بودن نتایج مشابه با سناریوهای یک یا دو پارامتری، ارائه نشده است. مطابق نتایج جدول ۱، هر چهار روش SVC، KNN، ANN و DT، در عمده سناریوها دارای درصد پیش‌بینی صحیح نزدیک به هم هستند. البته تعداد سناریوهای برتر در روش SVC نسبت به سه روش دیگر بیشتر است. در مجموع، بر اساس میزان درصد پیش‌بینی صحیح هر سناریو و همچنین سهولت و دقت اندازه‌گیری پارامترهای هواشناسی و جنبه‌های اقتصادی، می‌توان بهترین سناریوی پیش‌بینی شرایط بارش را برای هر روش و در بین چهار روش انتخاب نمود.

در روش SVC، سناریوی شماره ۸ با دو پارامتر حداکثر دما و درصد ابرناکی هوا با درصد پیش‌بینی صحیح برابر ۸۳/۵۶ بهترین عملکرد را نشان می‌دهد. البته سناریوهای تک‌پارامتری شماره ۲ و ۴، به ترتیب با پارامتر ورودی مجموع ساعات آفتابی و میانگین رطوبت نسبی با درصد پیش‌بینی صحیح تقریباً برابر ۸۲/۷۰، دارای عملکرد نزدیک به سناریوی ترکیبی شماره ۸ هستند. در روش KNN، بهترین عملکرد، مربوط به سناریوی شماره ۴ با پارامتر میانگین رطوبت نسبی هوا با درصد پیش‌بینی صحیح برابر ۸۲/۱۹ است. در روش ANN، سناریوی شماره ۷ با دو پارامتر ترکیبی میانگین رطوبت نسبی و درصد ابرناکی هوا با درصد پیش‌بینی صحیح برابر ۸۴/۱۰ و سناریوی شماره ۳ با پارامتر حداقل رطوبت نسبی با درصد پیش‌بینی صحیح برابر ۸۳/۴۴ به ترتیب دارای بهترین عملکرد در پیش‌بینی شرایط بارش روز بعد هستند. در روش DT، سناریوی شماره ۳ با پارامتر حداقل رطوبت نسبی با درصد پیش‌بینی صحیح برابر ۸۳/۸۳ بهترین سناریوی پیش‌بینی شرایط بارش روز بعد است. به‌طور کلی، نتایج سناریوهای مختلف هر چهار روش دارای عملکرد تقریباً مشابهی

در پیش‌بینی شرایط بارش روز بعد هستند. در بین انواع سناریوهای بررسی شده در چهار روش مختلف، به ترتیب سناریوی شماره ۷ روش ANN، سناریوی شماره ۳ روش DT و سناریوی شماره ۷ روش SVC بهترین عملکرد را دارند. همان‌طور که مشاهده می‌شود، سناریوی تک‌پارامتری شماره ۳ علاوه بر روش DT، در سه روش دیگر و در مقایسه با سایر سناریوها دارای عملکرد خوبی است (درصد پیش‌بینی صحیح برابر ۸۱/۰۵، ۸۱/۶۴ و ۸۳/۴۴ به ترتیب برای SVC، ANN و KNN). به نظر می‌رسد برای پیش‌بینی شرایط بارش یک روز بعد، با توجه به ملاحظات اقتصادی و ثبت لحظه‌ای پارامتر حداقل رطوبت نسبی، استفاده از این پارامتر می‌تواند مناسب‌ترین گزینه باشد.

جدول ۱: نتایج سناریوهای مختلف پیش‌بینی شرایط بارش ۱ روز بعد با استفاده از پارامترهای هواشناسی روزانه

درصد موارد پیش‌بینی صحیح (CCI)				پارامتر ورودی	شماره سناریو
DT	ANN	KNN	SVC		
۸۰/۶۴	۸۰/۸۲	۸۱/۰۹	۷۸/۰۸	Cloud ₁ ()	۱
۸۱/۳۶	۸۱/۹۲	۷۹/۶۶	۸۲/۷۰	Sun ₁ (hr)	۲
۸۳/۸۳	۸۳/۴۴	۸۱/۶۴	۸۱/۰۵	RHmin ₁ (%)	۳
۸۰/۴۱	۸۲/۰۶	۸۲/۱۹	۸۲/۷۳	RHmean ₁ (%)	۴
۸۰/۵۴	۷۶/۴۳	۸۰/۰۰	۷۶/۷۱	Tmax ₁ (°C)	۵
۸۱/۱۹	۸۲/۷۷	۸۰/۴۴	۸۳/۲۸	Cloud ₁ () and Sun ₁ (hr)	۶
۸۰/۶۰	۸۴/۱۰	۸۱/۷۸	۸۳/۰۱	Cloud ₁ () and RHmean ₁ (%)	۷
۸۱/۳۰	۸۲/۲۶	۸۰/۳۶	۸۳/۵۶	Cloud ₁ () and Tmax ₁ (°C)	۸
۸۰/۰۰	۸۱/۹۱	۷۹/۷۲	۸۲/۴۶	Sun ₁ (hr) and RHmin ₁ (%)	۹

منبع: یافته‌های پژوهش

در بین چهار روش مورد بررسی، تنها درخت تصمیم J48 دارای خروجی قابل استفاده برای مدیریت بارش و تصمیم‌گیری‌های آتی است. در ادامه، برای بهترین سناریوی مدل J48 (سناریوی شماره ۳)، نتایج درخت تصمیم‌گیری (رابطه ۱) و ماتریس همبستگی تعداد روزهای وقوع و عدم وقوع بارش بین داده‌های مشاهداتی و مدل درخت تصمیم (جدول ۲) ارائه شده است. بر اساس درخت تصمیم این سناریو در شهر اصفهان (رابطه ۱)، در روزهایی که حداقل رطوبت نسبی روزانه از مقدار ۲۸ درصد بیشتر شود، روز بعد بارش رخ خواهد داد. در غیر این صورت بارشی به وقوع نخواهد پیوست. بر اساس نتایج، میزان درستی این درخت تصمیم‌گیری در حدود ۸۴ درصد است؛ به عبارت دیگر، مطابق ماتریس همبستگی این درخت تصمیم (جدول ۲)، پیش‌بینی شرایط بارش روز بعد (وقوع یا عدم وقوع بارش) در ۳۰۷ روز یک سال یعنی تقریباً ۸۴ درصد موارد به صورت صحیح و در بقیه روزها، یعنی ۵۸ روز (۱۶ درصد)، به صورت ناصحیح انجام شده است؛ یعنی در مدل درخت تصمیم، ۳۳ روز در کلاس روزهای عدم وقوع بارش و ۲۵ روز در کلاس روزهای وقوع بارش، به اشتباه پیش‌بینی شده‌اند (جدول ۲).

$$RH_{min} \leq 28: N(288.0/33.0)$$

$$RH_{min} > 28: Y(77.0/25.0)$$

(۱)

جدول ۲: ماتریس همبستگی تعداد روزهای وقوع و عدم وقوع بارش بین داده‌های مشاهداتی و درخت تصمیم (سناریوی شماره ۳)

طبقه‌بندی روزها	روزهای عدم وقوع بارش (N) محاسباتی	روزهای وقوع بارش (Y) محاسباتی
روزهای وقوع بارش (Y) مشاهداتی	۳۳	۵۲
روزهای عدم وقوع بارش (N) مشاهداتی	۲۵۵	۲۵

منبع: یافته‌های پژوهش

در جدول ۳، نتایج سناریوهای مختلف پیش‌بینی شرایط بارش ۲ روز بعد با استفاده از پارامترهای هواشناسی روزانه ارائه شده است. مطابق نتایج این جدول، دو سناریوی تک‌پارامتری با پارامترهای حداکثر رطوبت نسبی و میانگین رطوبت نسبی، دارای بیشترین درصد موارد پیش‌بینی صحیح (CCI) هستند. مقادیر درصد پیش‌بینی صحیح این دو سناریو در هر چهار روش مورد بررسی، تقریباً مشابه هم هستند. مقایسه نتایج هر سناریو در بین روش‌های مورد بررسی نیز نشان می‌دهد که روش DT دارای بیشترین مقدار درصد پیش‌بینی صحیح است. عملکرد سه روش دیگر نیز مشابه بوده و تقریباً مزیت خاصی نسبت به هم ندارند. به نظر می‌رسد برای پیش‌بینی شرایط بارش ۲ روز بعد، سناریوی شماره ۱ با پارامتر ورودی حداکثر رطوبت نسبی با توجه به ثبت لحظه‌ای مقدار آن، می‌تواند به‌عنوان مناسب‌ترین مدل تصمیم‌گیری انتخاب گردد.

در رابطه ۲ و جدول ۴، به ترتیب نتایج درخت تصمیم‌گیری و ماتریس همبستگی تعداد روزهای وقوع و عدم وقوع بارش بین داده‌های مشاهداتی و مدل درخت تصمیم سناریوی ۱ ارائه شده است. بر اساس درخت تصمیم این سناریو در شهر اصفهان (رابطه ۲)، در روزهایی که حداکثر رطوبت نسبی روزانه از مقدار ۶۱ درصد بیشتر شود، ۲ روز بعد بارش رخ خواهد داد، در غیر این صورت بارشی به وقوع نخواهد پیوست. بر اساس نتایج، میزان درستی این تصمیم‌گیری ۷۹/۱۲ درصد است. نتایج ماتریس همبستگی این درخت تصمیم (جدول ۴) نیز نشان می‌دهد که پیش‌بینی شرایط بارش ۲ روز بعد (وقوع یا عدم وقوع بارش) در ۲۸۸ روز یک سال یعنی تقریباً ۷۹/۱۲ درصد موارد به‌صورت صحیح و در بقیه روزها یعنی ۷۶ روز (۲۰/۸۳ درصد) به‌صورت ناصحیح انجام شده است؛ به‌عبارت‌دیگر، در مدل درخت تصمیم، ۲۳ روز در کلاس روزهای عدم وقوع بارش و ۵۳ روز در کلاس روزهای وقوع بارش، به اشتباه پیش‌بینی شده‌اند.

جدول ۳: نتایج سناریوهای مختلف پیش‌بینی وقوع بارش ۲ روز بعد با استفاده از پارامترهای هواشناسی روزانه

درصد موارد پیش‌بینی صحیح (CCI)				پارامتر ورودی	شماره سناریو
DT	ANN	KNN	SVC		
۷۹/۱۲	۷۷/۳۷	۷۶/۶۴	۷۶/۹۲	RHmax ₂ (%)	۱
۸۰/۲۱	۷۹/۳۹	۷۷/۷۴	۷۶/۳۷	RHmean ₂ (%)	۲

منبع: یافته‌های پژوهش

$$RH_{max} \leq 61: N(250.0/23.0)$$

$$RH_{max} > 61: Y(114.0/53.0)$$

(۲)

جدول ۴: ماتریس همبستگی تعداد روزهای وقوع و عدم وقوع بارش بین داده‌های مشاهداتی و درخت تصمیم (سناریوی شماره ۱)

طبقه‌بندی روزها	روزهای عدم وقوع بارش (N) محاسباتی	روزهای وقوع بارش (Y) محاسباتی
روزهای وقوع بارش (Y) مشاهداتی	۲۳	۶۱
روزهای عدم وقوع بارش (N) مشاهداتی	۲۲۷	۵۳

منبع: یافته‌های پژوهش

در حالت تأخیر ۳ روزه، بررسی عملکرد سناریوهای مختلف در چهار روش نشان داد که اکثر سناریوها دارای درصد پیش‌بینی صحیح کمتر از ۷۷ درصد هستند. مطابق جدول ۵، تنها یک سناریو با پارامتر ورودی میانگین رطوبت نسبی در هر چهار روش، دارای درصد موارد پیش‌بینی صحیح (CCI) بالای ۷۷ درصد بود. در بین روش‌های مورد بررسی نیز روش DT دارای بهترین عملکرد پیش‌بینی صحیح بود (۸۱/۵۴ درصد). در رابطه ۳ و جدول ۶ به ترتیب نتایج درخت تصمیم‌گیری و ماتریس همبستگی تعداد روزهای وقوع و عدم وقوع بارش بین داده‌های مشاهداتی و مدل درخت تصمیم سناریوی ۱ ارائه شده است. بر اساس درخت تصمیم این سناریو در شهر اصفهان (رابطه ۳)، در روزهایی که میانگین رطوبت نسبی روزانه از مقدار تقریبی ۴۴/۲۷ درصد بیشتر شود، ۳ روز بعد بارش رخ خواهد داد، در غیر این صورت بارشی به وقوع نخواهد پیوست. بر اساس نتایج، میزان درستی این تصمیم‌گیری در حدود ۸۱/۵۴ درصد است. نتایج ماتریس همبستگی این درخت تصمیم (جدول ۶) نشان می‌دهد که پیش‌بینی شرایط بارش ۳ روز بعد (وقوع یا عدم وقوع بارش) در ۲۹۶ روز یک سال یعنی تقریباً ۸۱/۵۴ درصد موارد به‌صورت صحیح و در بقیه روزها یعنی ۶۷ روز (۱۸/۴۶ درصد) به‌صورت ناصحیح انجام شده است؛ یعنی در مدل درخت تصمیم، ۲۸ روز در کلاس روزهای عدم وقوع بارش و ۳۹ روز در کلاس روزهای وقوع بارش، به اشتباه پیش‌بینی شده‌اند.

جدول ۵: نتایج سناریوهای مختلف پیش‌بینی وقوع بارش ۳ روز بعد با استفاده از پارامترهای هواشناسی روزانه

شماره سناریو	پارامتر ورودی	درصد موارد پیش‌بینی صحیح (CCI)			
		DT	ANN	KNN	SVC
۱	RHmean ₃ (%)	۸۱/۵۴	۸۰/۷۱	۸۱/۲۶	۷۷/۶۸

منبع: یافته‌های پژوهش

$$RH_{mean} \leq 44.27: N(269.0/28.0)$$

(۳)

$$RH_{mean} > 44.27: Y(94.0/39.0)$$

جدول ۶: ماتریس همبستگی تعداد روزهای وقوع و عدم وقوع بارش بین داده‌های مشاهداتی و درخت تصمیم (سناریوی شماره ۱)

طبقه‌بندی روزها	روزهای عدم وقوع بارش (N) محاسباتی	روزهای وقوع بارش (Y) محاسباتی
روزهای وقوع بارش (Y) مشاهداتی	۲۸	۵۵
روزهای عدم وقوع بارش (N) مشاهداتی	۲۴۱	۳۹

منبع: یافته‌های پژوهش

در جدول ۷، نتایج سناریوهای مختلف پیش‌بینی شرایط بارش ۴ روز بعد با استفاده از پارامترهای هواشناسی روزانه ارائه شده است. در بین سناریوهای تک‌پارامتری، چهار سناریو به ترتیب با پارامترهای ورودی میانگین رطوبت نسبی و حداکثر، میانگین و حداقل دمای هوا دارای بیشترین درصد موارد پیش‌بینی صحیح (CCI) بودند. در ترکیب پارامترها، مشابه نتایج قبلی، ترکیب‌های دو پارامتره، بهترین عملکرد را داشتند و افزایش تعداد پارامترها، تأثیری در افزایش عملکرد سناریوهای ترکیبی نداشت. در دو روش SVC و DT، بهترین عملکرد مربوط به سناریوی شماره ۵ با ترکیب دو پارامتر حداکثر دما و ساعات آفتابی با درصد پیش‌بینی صحیح به ترتیب برابر ۷۷/۳۴ و ۸۰/۶۶ درصد است. در دو روش KNN و ANN بهترین عملکرد مربوط به سناریوهای ترکیبی شماره ۶ با پارامترهای ورودی ساعات آفتابی و حداقل دمای هوا با درصد پیش‌بینی صحیح به ترتیب برابر ۸۰/۳۵ و ۸۱/۲۱ درصد است. همان‌طور که مشاهده می‌شود، تعداد سناریوهای برتر مربوط به روش درخت تصمیم است. در این روش، سناریوی شماره ۲ با پارامتر حداکثر دمای هوا و مقدار CCI برابر ۸۰/۳۸ درصد، دارای عملکرد قابل قبول در مقایسه با سناریوی ترکیبی شماره ۵ یا ۶ (بهترین سناریوها) است. به نظر می‌رسد برای پیش‌بینی شرایط بارش ۴ روز بعد، با توجه به ملاحظات اقتصادی و ثبت لحظه‌ای حداکثر دمای هوا، سناریوی شماره ۲ می‌تواند به‌عنوان گزینه کاربردی انتخاب شود. شایان‌ذکر است این سناریو در سه روش دیگر دارای عملکرد قابل قبول نسبت به سایر سناریوها است.

در رابطه ۴ و جدول ۸، به ترتیب نتایج الگوریتم درخت تصمیم‌گیری سناریوی شماره ۲ و ماتریس همبستگی تعداد روزهای وقوع و عدم وقوع بارش بین داده‌های مشاهداتی و مدل درخت تصمیم این سناریو ارائه شده است. بر اساس درخت تصمیم این سناریو در شهر اصفهان (رابطه ۴)، در روزهایی که حداکثر دمای هوای روزانه از مقدار تقریبی ۱۳ درجه کمتر شود، ۴ روز بعد بارش رخ خواهد داد. در غیر این صورت بارشی به وقوع نخواهد پیوست. بر اساس نتایج، میزان درستی این تصمیم‌گیری ۸۰/۳۸ درصد است. نتایج ماتریس همبستگی این درخت تصمیم (جدول ۸) نشان می‌دهد که پیش‌بینی شرایط بارش روز بعد (وقوع یا عدم وقوع بارش) در ۲۹۱ روز یک سال یعنی تقریباً ۸۰/۳۸ درصد به‌صورت صحیح و در بقیه روزها یعنی ۷۱ روز (۱۹/۶۲ درصد) به‌صورت ناصحیح انجام شده است؛ به‌عبارت‌دیگر، در مدل درخت تصمیم، ۴۵ روز در کلاس روزهای عدم وقوع بارش و ۲۶ روز در کلاس روزهای وقوع بارش، به اشتباه پیش‌بینی شده‌اند.

جدول ۷: نتایج سناریوهای مختلف پیش‌بینی وقوع بارش ۴ روز بعد با استفاده از پارامترهای هواشناسی روزانه

درصد موارد پیش‌بینی صحیح (CCI)				پارامتر ورودی	شماره سناریو
DT	ANN	KNN	SVC		
۷۹/۸۳	۸۰/۶۰	۸۰/۰۰	۷۶/۳۴	RHmean ₄ (%)	۱
۸۰/۳۸	۷۹/۳۵	۷۹/۲۸	۷۷/۰۷	Tmax ₄ (°C)	۲
۸۰/۱۱	۷۹/۸۳	۷۸/۱۷	۷۷/۰۴	Tmean ₄ (°C)	۳
۷۹/۰۰	۷۷/۶۴	۷۹/۵۵	۷۷/۰۷	Tmin ₄ (°C)	۴
۸۰/۶۶	۸۰/۳۰	۷۸/۱۷	۷۷/۳۴	Tmax ₄ (°C) and Sun ₄ (hr)	۵
۷۹/۷۵	۸۱/۲۱	۸۰/۳۵	۷۶/۵۱	Tmin ₄ (°C) and Sun ₄ (hr)	۶

منبع: یافته‌های پژوهش

$$T_{\max} \leq 12.93:Y(64.0/26.0) \quad (4)$$

$$T_{\max} > 12.93:N(298.0/45.0)$$

جدول ۸: ماتریس همبستگی تعداد روزهای وقوع و عدم وقوع بارش بین داده‌های مشاهداتی و درخت تصمیم (سناریوی شماره ۲)

طبقه‌بندی روزها	روزهای عدم وقوع بارش (N) محاسباتی	روزهای وقوع بارش (Y) محاسباتی
روزهای وقوع بارش (Y) مشاهداتی	۴۵	۳۸
روزهای عدم وقوع بارش (N) مشاهداتی	۲۵۳	۲۶

منبع: یافته‌های پژوهش

در جدول ۹، نتایج سناریوهای مختلف پیش‌بینی شرایط بارش ۵ روز بعد با استفاده از پارامترهای هواشناسی روزانه ارائه شده است. از بین سناریوهای تک‌پارامتری، فقط دو سناریو با پارامترهای ورودی متوسط دمای هوا و سرعت باد، دارای عملکرد بالایی بودند. در هر چهار روش، تفاوت قابل توجهی بین عملکرد سناریوهای تک‌پارامتری و دو پارامتری وجود ندارد. در روش SVC، بهترین عملکرد مربوط به سناریوی ترکیبی شماره ۴ با پارامترهای ورودی حداکثر رطوبت نسبی و ساعات آفتابی با درصد موارد پیش‌بینی صحیح (CCI) برابر ۷۸/۴۷ درصد است. در روش KNN، بهترین عملکرد مربوط به سناریوی تک‌پارامتری شماره ۱ (میانگین دمای هوا) با مقدار درصد پیش‌بینی صحیح ۸۰/۷۲ است. در روش ANN، بهترین عملکرد مربوط به سناریوی ترکیبی شماره ۲ با پارامتر ورودی سرعت باد با مقدار درصد پیش‌بینی صحیح ۷۹/۲۲ درصد است. در روش DT، بهترین عملکرد مربوط به سناریوی ترکیبی شماره ۴ با پارامترهای ورودی ساعات آفتابی و سرعت باد است. این سناریو با عملکرد پیش‌بینی صحیح برابر ۸۰/۶۰ درصد، دارای بالاترین عملکرد در بین انواع سناریو است. در این روش، علاوه بر این سناریو، سناریوی شماره ۱ (میانگین دمای هوا) و سناریوی شماره ۲ (سرعت باد) به ترتیب با درصد پیش‌بینی صحیح برابر ۸۰/۳۳ و ۸۰/۰۵ دارای عملکرد مناسبی هستند. به نظر می‌رسد برای پیش‌بینی شرایط بارش ۵ روز بعد، با توجه به ملاحظات اقتصادی، دقت و سهولت ثبت پارامتر، سناریوی شماره ۱ می‌تواند به‌عنوان گزینه کاربردی انتخاب گردد.

در رابطه ۵ و جدول ۱۰، به ترتیب نتایج درخت تصمیم‌گیری و ماتریس همبستگی تعداد روزهای وقوع و عدم وقوع بارش بین داده‌های مشاهداتی و مدل درخت تصمیم سناریوی شماره ۱ ارائه شده است. بر اساس درخت تصمیم این سناریو در شهر اصفهان (رابطه ۵)، در روزهایی که میانگین دمای روزانه از مقدار ۷/۸۷ درجه کمتر باشد، ۵ روز بعد بارش رخ خواهد داد. در غیر این صورت بارشی به وقوع نخواهد پیوست. بر اساس نتایج، میزان درستی این تصمیم‌گیری در حدود ۸۰/۳۳ درصد است. نتایج ماتریس همبستگی این درخت تصمیم (جدول ۱۰) نشان می‌دهد که پیش‌بینی شرایط بارش ۵ روز بعد (وقوع یا عدم وقوع بارش) در ۲۸۹ روز یک سال یعنی ۸۰/۳۳ درصد به‌صورت صحیح و در بقیه روزها یعنی ۷۲ روز (۱۹/۶۷ درصد) به‌صورت ناصحیح انجام شده است؛ به‌عبارت‌دیگر، در مدل درخت تصمیم، ۳۷ روز در کلاس روزهای عدم وقوع بارش و ۳۵ روز در کلاس روزهای وقوع بارش، به اشتباه پیش‌بینی شده‌اند.

جدول ۹: نتایج سناریوهای مختلف پیش‌بینی وقوع بارش ۵ روز بعد با استفاده از پارامترهای هواشناسی روزانه

درصد موارد پیش‌بینی صحیح (CCI)				پارامتر ورودی	شماره سناریو
DT	ANN	KNN	SVC		
۸۰/۳۳	۷۸/۱۱	<u>۸۰/۷۲</u>	۷۷/۰۰	Tmean ₅	۱
<u>۸۰/۰۵</u>	<u>۷۹/۲۲</u>	۷۹/۵۰	۷۷/۰۰	Wind ₅	۲
<u>۷۹/۷۷</u>	۷۸/۳۹	۷۷/۸۳	۷۸/۶۷	RHmean ₅ (%) and Sun ₋₅ (hr)	۳
<u>۸۰/۶۰</u>	۷۸/۹۴	۸۰/۰۵	<u>۷۸/۴۷</u>	Wind ₅ (m/s) and Sun ₋₅ (hr)	۴
<u>۷۹/۱۵</u>	۷۸/۶۲	۸۰/۲۳	۷۶/۱۷	RHmax ₅ (%), Sun ₋₅ (hr)	۵

منبع: یافته‌های پژوهش

$$T_{\text{mean}} \leq 7.87:Y(79.0/35.0) \quad (5)$$

$$T_{\text{mean}} > 7.87:N(282.0/37.0)$$

جدول ۱۰: ماتریس همبستگی تعداد روزهای وقوع و عدم وقوع بارش بین داده‌های مشاهداتی و درخت تصمیم (سناریوی شماره ۱)

طبقه‌بندی روزها	روزهای عدم وقوع بارش (N) محاسباتی	روزهای وقوع بارش (Y) محاسباتی
روزهای وقوع بارش (Y) مشاهداتی	۳۷	۴۴
روزهای عدم وقوع بارش (N) مشاهداتی	۲۴۵	۳۵

منبع: یافته‌های پژوهش

در جدول ۱۱، نتایج صحت‌سنجی بهترین سناریوهای پیش‌بینی شرایط وقوع بارش روزانه با استفاده از داده‌های هواشناسی دوره ۲۰۱۶-۲۰۱۱ ارائه شده است. مطابق نتایج، عملکرد بهترین سناریوی پیش‌بینی شرایط وقوع بارش ۱ روز بعد (با پارامتر ورودی حداقل رطوبت نسبی) در دوره صحت‌سنجی تقریباً مشابه عملکرد این سناریو در دوره آماری ۲۰۰۹-۲۰۰۰ می‌باشد. برای پیش‌بینی شرایط وقوع بارش ۲ و ۳ روز بعد، نتایج بخش صحت‌سنجی نشان‌دهنده کاهش ۸/۵ درصدی عملکرد بهترین سناریوها است. عملکرد بهترین سناریوی پیش‌بینی شرایط وقوع بارش ۴ روز بعد (با پارامتر ورودی دمای حداکثر) در دوره صحت‌سنجی تقریباً مشابه عملکرد این سناریو در دوره آماری ۲۰۰۹-۲۰۰۰ می‌باشد. در نهایت برای پیش‌بینی شرایط وقوع بارش ۵ روز بعد، نتایج بخش صحت‌سنجی نشان‌دهنده کاهش ۵ درصدی عملکرد بهترین سناریوها است. در مجموع مطابق این نتایج می‌توان گفت، با استفاده از داده‌های روزانه حداقل رطوبت نسبی و یا دمای حداکثر، به ترتیب شرایط وقوع بارش ۱ و ۴ روز بعد با دقت مناسبی قابل پیش‌بینی است.

جدول ۱۱: نتایج صحت‌سنجی بهترین سناریوهای پیش‌بینی شرایط بارش روزانه با استفاده از داده‌های ۲۰۱۶-۲۰۱۱

پیش‌بینی شرایط بارش	پارامتر ورودی	درصد پیش‌بینی صحیح	درصد پیش‌بینی ناصحیح
یک روز بعد	RHmin ₁	۸۲/۶۹	۱۷/۳۱
۲ روز بعد	RHmax ₂	۷۰/۴۷	۲۹/۵۳
۳ روز بعد	RHmean ₃	۷۲/۹۶	۲۷/۰۴
۴ روز بعد	Tmax ₄	۷۸/۶۸	۲۱/۳۲
۵ روز بعد	Tmean ₅	۷۵/۱۲	۲۴/۸۸

منبع: یافته‌های پژوهش

بحث و نتیجه‌گیری

در این مطالعه، پیش‌بینی شرایط بارش روزانه در شهر اصفهان به کمک چهار مدل هوشمند SVC، KNN، ANN و DT و با استفاده از داده‌های روزهای قبل (۱ تا ۷ روز) انجام گردید. مقایسه نتایج تأخیرهای ۱ تا ۵ روزه نشان می‌دهد که با استفاده از داده‌های هواشناسی روزانه می‌توان شرایط بارش تا ۵ روز بعد را با دقت خوبی (۸۰ درصد) پیش‌بینی نمود. دقت پیش‌بینی وقوع بارش با استفاده از داده‌های مربوط به ۶ یا ۷ روز قبل نسبت به داده‌های ۱ الی ۵ روز قبل، پایین بود. به‌طور کلی، با نزدیک شدن به روز مورد نظر، استفاده از پارامتر رطوبت نسبی منجر به پیش‌بینی بهتر شرایط بارش روزانه شد. عملکرد هر چهار مدل هوشمند در پیش‌بینی شرایط وقوع بارش در انواع سناریوها تقریباً نزدیک به هم بود. نتایج این پژوهش یا یافته‌های مطالعه ویراسینگ و همکاران (۲۰۱۰) مطابقت خوبی دارد که در مطالعه خود گزارش کردند مدل شبکه عصبی مصنوعی (ANN) می‌تواند شرایط وقوع بارش روزانه را با دقت ۷۹ درصد پیش‌بینی کند. همچنین برخی پژوهشگران از جمله بهاتاچاریا و سولوماتین (۲۰۰۵)، مدرسی و عراقی‌نژاد (۲۰۱۴) و ستاری و همکاران (۲۰۱۶) در مطالعات خود بر توانایی و عملکرد مشابه روش‌های مختلف داده‌کاوی از جمله روش‌های بردار پشتیبان، K نزدیک‌ترین همسایگی، مدل درختی و شبکه عصبی مصنوعی در پیش‌بینی پارامترهای هواشناسی و هیدرولوژیکی اشاره نموده‌اند. در مدل درخت تصمیم، به علت ارائه درخت تصمیم‌گیری ساده و کاربردی، نتایج بهترین سناریوهای این روش در تأخیرهای ۱ تا ۵ روزه ارائه گردید. در شرایط تأخیرهای ۱، ۲ و ۳ روزه، مناسب‌ترین پارامتر برای پیش‌بینی شرایط بارش روزانه، به ترتیب حداقل، حداکثر و میانگین رطوبت نسبی هوا است. در حالی که در تأخیر ۴ و ۵ روزه، بهترین پارامتر به ترتیب حداکثر و میانگین دمای هوا است. در شرایط تأخیر ۱ تا ۳ روزه و نزدیک شدن به‌روز مورد نظر، رطوبت هوا به‌عنوان پارامتر غالب در پیش‌بینی شرایط بارش ظاهر شده است. با توجه این که رطوبت هوا از عوامل تأثیرگذار در فرایند بارش است، به نظر می‌رسد این بخش از نتایج، همخوانی خوبی با فیزیک حاکم بر پدیده بارش دارد. همچنین نتایج این بخش از پژوهش مطابقت خوبی با نتایج پژوهش‌های بهرامی و همکاران (۱۳۹۶) و آنیهوتوری و موهاپاترا (۲۰۱۲) دارد. بهرامی و همکاران (۱۳۹۶) با بررسی پیش‌بینی بارش ماهانه ایستگاه آباد به کمک شبکه عصبی مصنوعی گزارش کردند که دو پارامتر رطوبت نسبی و حداکثر دما بیشترین تأثیر را در پیش‌بینی بارش دارند. آنیهوتوری و موهاپاترا (۲۰۱۲) نیز با مطالعه پیش‌بینی وقوع باران‌های موسمی روزانه تابستان در منطقه کارناتاکای هند نتیجه گرفتند که بیشترین همبستگی وقوع بارش روزانه با پارامترهای رطوبت نسبی و حداقل دمای هوا روزانه است. به‌طور کلی در مدیریت پدیده‌های هواشناسی و هیدرولوژیکی، نقش زمان و داشتن فرصت مناسب برای برنامه‌ریزی و انجام اقدامات لازم، پارامتر بسیار کلیدی است. از این رو پیش‌بینی وقوع پدیده‌های هواشناسی، خصوصاً بارش و شرایط آن‌ها در آینده می‌تواند کمک بسزایی در مدیریت این پدیده داشته باشد. امروزه با توسعه مدل‌های هوشمند کاوش داده‌ها، پیش‌بینی شرایط هواشناسی با استفاده از داده‌های مربوط به گذشته‌های دور و یا نزدیک، امری ممکن و سهل به نظر می‌رسد. در این میان، وقوع بارش، یکی از مهم‌ترین پدیده‌های هواشناسی، چنانچه با استفاده از داده‌های هواشناسی روز یا روزهای قبل پیش‌بینی گردد، می‌تواند کمک شایانی در مدیریت و برنامه‌ریزی انواع فعالیت‌های کشاورزی، گردشگری و غیره کند.

منابع

- ۱- بابا علی، حمیدرضا و رضا دهقانی. (۱۳۹۶): مقایسه مدل‌های هوشمند در تخمین بارش ماهانه حوضه کاکا رضا. اکوهیدرولوژی، دوره چهارم، شماره اول، صفحات ۱۱-۱، تهران.
- ۲- بهرامی، مهدی. محمدجواد امیری. فاطمه رضایی مهار لویی و کرامت ... غفاری. (۱۳۹۶): تعیین اثر پیش‌پردازش داده بر عملکرد شبکه عصبی مصنوعی به‌منظور پیش‌بینی بارش ماهانه در شهرستان آباده. اکوهیدرولوژی، دوره چهارم، شماره اول، صفحات ۳۷-۲۹، تهران.
- ۳- ستاری، محمدتقی. علی رضازاده جودی و فرناز نهرین. (۱۳۹۳): پیش‌بینی مقادیر بارش ماهانه با استفاده از شبکه‌های عصبی مصنوعی و مدل درختی M5 (مطالعه موردی: ایستگاه اهر). پژوهش‌های جغرافیای طبیعی، دوره چهل و ششم، شماره دوم، صفحات ۲۶۰-۲۴۷، تهران.
- ۴- فلاحی، محمدرضا. هادی وروانی و سعید گلیان (۱۳۹۰): پیش‌بینی بارش با استفاده از مدل رگرسیون درختی به‌منظور کنترل سیل، پنجمین کنفرانس سراسری آب‌خیزداری و مدیریت منابع آب‌و خاک کشور، دانشگاه شهید باهنر، کرمان.
- ۵- فقیه، همایون. (۱۳۸۹): ارزیابی کاربرد شبکه عصبی مصنوعی و بهینه‌سازی آن با روش الگوریتم ژنتیک در تخمین داده‌های بارش ماهانه (مطالعه موردی: منطقه کردستان). علوم و فنون کشاورزی و منابع طبیعی (علوم آب‌و خاک)، دوره چهاردهم، شماره پنجاه و یکم، صفحات ۴۲-۲۷، اصفهان.
- 6- Agnihotri, G. Mohapatra, M. (2012): Prediction of Occurrence of Daily Summer Monsoon Precipitation Over Karnataka. *Meteorological Applications*, Vol. 19, PP. 130-139.
- 7- Bhattacharya, B. Solomatine, D.P. (2005): Neural Networks and M5 Model Trees in Modeling Water Level-Discharge Relationship. *Neuro Computing*, Vol. 63, PP. 381-396.
- 8- Freiwan, M. Cigizoglu, H.K. (2005): Prediction of Total Monthly Rainfall in Jordan Using Feed Forward Backpropagation Method. *Fresenius Environmental Bulletin*, Vol. 14 No. 2, PP. 142-151.
- 9- George, J. Letha, J. Jairaj, P.G. (2016): Daily Rainfall Prediction Using Generalized Linear Bivariate Model - A Case Study. *Procedia Technology*, Vol. 24, PP. 31-38.
- 10- Mandal, S. Choudhury, B.U. (2014): Estimation and Prediction of Maximum Daily Rainfall at Sagar Island Using Best Fit Probability Models. *Theoretical Applied Climatology*, Vol. 121, No. 1-2, PP. 1-11.
- 11- Mekanik, F. Imteaz, M.A. Gato-Trinida, S. Elmahdi, A. (2013): Multiple Regression and Artificial Neural Network For Long Term Rainfall Forecasting Using Large Scale Climate Modes. *Journal of Hydrology*, Vol. 503, PP. 11-21.
- 12- Modaresi, F. Araghinejad, S. (2014); A Comparative Assessment of Support Vector Machines, Probabilistic Neural Networks and K-Nearest Neighbors' Algorithms for Water Quality Classification. *Water Resources Management*, Vol. 28, PP. 4095-4111.
- 13- Moghaddamnia, A. Ghafari Gousheh, M. Piri J. Amin S. Han, D. (2009): Evaporation Estimation Using Artificial Neural Networks and Adaptive Neurofuzzy Inference System Techniques. *Advances in Water Resources*, Vol. 32, PP. 88-97.
- 14- Pal, M. Mather, P.M. (2003): An Assessment of The Effectiveness of Decision Tree Methods for Land Cover Classification. *Remote Sensing Environment*, Vol. 86, No. 4, PP. 554-565.
- 15- Partal, T. Cigizoglu H.K. (2009): Prediction of Daily Precipitation Using Wavelet-Neural Networks. *Hydrological Sciences Journal*, Vol. 54, No. 2, PP. 234-246.
- 16- Saghebian, S.M. Sattari, M.T. Mirabbasi, R. Pal, M. (2013): Ground Water Quality Classification By Decision Tree Method in Ardebil Region, Iran. *Arabian Journal of Geosciences*, Vol. 7, No. 11, PP. 4767-4777.
- 17- Sattari, M.T. Rezazadeh Joudi, A. Kusiak, A. (2016): Estimation of Water Quality Parameters With Data-Driven Model. *American Water Works Association*, Vol. 108, No. 4, PP. 232-239.

- 18- Vapnik, V. (1995): *The Nature of Statistical Learning Theory*. New York, Springer-Verlag.
- 19- Weerasinghe, H.D.P. Premaratne, H.L. Sonnadara, D.U.J. (2010): Performance of Neural Networks in Forecasting Daily Precipitation Using Multiple Sources. *Journal of The National Science Foundation of Sri Lanka*, Vol. 38, No. 3, PP. 163-170.
- 20- Witten, I.H. Frank, E. Hall, M.A. (2005): *Data Mining: Practical Machine Learning Tools and Techniques*. California, Morgan Kaufmann.
- 21- Wu, C.L. Chau, K. W. Fan, C. (2010): Prediction of Rainfall Time Series Using Modular Artificial Neural Networks Coupled With Data Preprocessing Techniques. *Journal of Hydrology*, Vol. 389, No. 1-2, PP. 146-167.
- 22- Yang Su, M. (2011): Real-Time Anomaly Detection Systems for Denial-of-Service Attacks By Weighted K-Nearest Neighbor Classifiers. *Expert Systems With Applications*, Vol. 38, No. 4, PP. 3492-3498.