دانشگاه علوم کشاورزی و منابع طبیعی ساری
پژوهشنامه مدیریت حوزه آبخیز

بررسی تأثیر تعدیل دامنه تغییرات داده ها بر کارایی مدل شبکه عصبی مصنوعی و درخت تصمیم رگرسیونی در پیش بینی خشکسالی

ا. حبیبی پور[۱]، م. ت. دستورانی[۲]، م. ر. اختصاصی[۲] و ح. افخمی[۳]

۱- کارشناس ارشد آبخیزداری دانشگاه یزد، نویسنده مسوول: habibipoor_azam @yahoo.com
۲- دانشیار دانشکده منابع طبیعی دانشگاه یزد
۳- کارشناس ارشد آبخیزداری دانشگاه یزد

چکیده

خشکسالی یکی از اثرات تغییر سامانه اقلیمی است. پیش بینی خشکسالی نقش مهمی در اعمال روش های مؤثر مدیریت منابع آب ایفا می کند. روش های مختلفی برای مطالعه خشکسالی وجود دارد. روش تحلیل داده های بارندگی، جزء عمومی روش‌های تحلیل خشکسالی به شمار می‌رود. لذا پیش بینی دقیق و پیش از وقوع بارش می تواند شرایط را برای ارزیابی وضعیت خشکسالی فراهم نماید. هدف از این پژوهش، بررسی تأثیر پیش پردازش داده ها بر عملکرد دو مدل داده کاوی در پیش بینی خشکسالی در ایستگاه سینوپتیک یزد می باشد. در این رابطه از دو روش شبکه عصبی مصنوعی و درخت تصمیم رگرسیونی که از انواع روش های داده کاوی محسوب می شود استفاده شد و شبیه سازی ها در دو حالت کلی صورت گرفت. در حالت اول، از مقادیر اصلی برخی پارامترهای اقلیمی استفاده و میزان بارش ۱۲ ماه پیش از وقوع پیش بینی گردید. در حالت دوم میانگین متحرک سه ساله همان داده ها به مدل معرفی و پیش بینی بر همین اساس انجام شد. در پایان برای ارزیابی دقت و درستی دو روش مورد استفاده، معیارهای آماری RMSE، R مورد استفاده قرار گرفت. یافته ها نشان داد که اعمال میانگین لغزان روی داده های اصلی به نحو چشمگیری در بهبود کارایی هر دو مدل مؤثر می باشد و در این شرایط هر دو روش درخت تصمیم رگرسیونی و شبکه عصبی مصنوعی در ایستگاه یزد قادرند با ضریب اطمینان بالایی میزان بارش را ۱۲ ماه پیش از وقوع برآورد نمایند.

واژه های کلیدی: پیش بینی، خشکسالی، شبکه عصبی مصنوعی، درخت تصمیم رگرسیونی، داده کاوی

مقدمه

در دهه های اخیر تغییر سیستم اقلیمی زمین به ویژه گرمایش سیاره ای ناشی از مصرف سوخت های فسیلی به طور گسترده ای مورد توجه قرار گرفته است. هیئت بین الدول تغییر اقلیم در سال ۲۰۰۱ گزارش داد که اقلیم در حال تغییر و گرمایش جهانی در حال وقوع است. توجه به تغییرات اقلیمی در سال های اخیر به علت نتایج اقتصادی، اجتماعی و

[Downloaded from jwmr.sanru.ac.ir on 2026-10-10]

آسیب های مالی مربوط به رویدادهای جوی اهمیت زیادی پیدا کرده است. یکی از اثرات تغییر سیستم اقلیمی کره زمین بروز خشکسالی در بسیاری از نقاط کره زمین و یا تشدید سیل و سیل گیری در نقاط دیگر است که برخلاف سایر بلایای طبیعی همانند طوفان، آتشفشان، زمین لرزه، سونامی و غیره به آرامی آغاز و به طور خزنده ای گسترش می یابد و تقریباً در همه رژیم های اقلیمی دیده می شود (۱۴).

خشکسالی در زمینه های مختلف، هم به طور مستقیم و هم به طور غیر مستقیم اثرهای زیان آوری وارد می کند. با کم شدن بارندگی یا نبود آن برای یک مدت طولانی در سال (خشکسالی)، مراتع، جنگل ها، مزارع و باغ هایی که منبع آبی آن ها وابسته به ریزش های جوی است و نیز خاک و دیگر منابع طبیعی به طور مستقیم زیان می بینند. علاوه بر آن منابع آب های سطحی و زیرزمینی نیز از ریزش های جوی متأثر می گردند. به این ترتیب با وقوع خشکسالی، فعالیت ها و تأسیسات وابسته به آب، چه شهری و روستایی، چه صنعتی و غیره تحت تأثیر قرار گرفته و خسارت خواهند دید (۱۰).

از این رو پیش بینی خشکسالی می تواند در اعمال روشهای مؤثر مدیریت منابع آب و تعدیل خسارات ناشی از خشکسالی مؤثر واقع شود. روش های مختلفی برای مطالعه خشکسالی وجود دارد. روش تحلیل داده های بارندگی، جزء عمومی روش‌های تحلیل خشکسالی به شمار می‌رود. علت این امر دسترسی راحت‌تر و آسان‌تر به انواع داده‌های بارندگی در بخش‌های مختلف کره زمین است. از طرفی مقادیر بارندگی جزء بی‌ثبات‌ترین متغیرهای آب و هوایی مخصوصاً در مناطق خشک محسوب می‌شود و از این رو نمایه خوبی برای مطالعه خشکسالی می‌باشد (۴). این در حالی است که در سال های اخیر شاهد حرکتی مستمر، از تحقیقات صرفاً تئوری به پژوهش های کاربردی به خصوص در زمینه پردازش اطلاعات، برای مسائلی که برای آنها راه حلی موجود نیست و یا به راحتی قابل حل نیستند، بوده ایم (۱). این موضوع تحت عنوان کشف دانش و داده کاوی تعبیر شده است. داده کاوی یک حوزه میان رشته ای و با رشد سریع است که حوزه های مختلفی همچون پایگاه داده، آمار، یادگیری ماشین و سایر زمینه های مرتبط را تلفیق کرده است تا اطلاعات و دانش ارزشمند نهفته در حجم بزرگی از داده‌ها را استخراج نماید (۱۱).

امروزه حجم زیادی از داده‌ها توسط سازمان ها و ارگان های متولی امر در مقیاس های زمانی مختلف ثبت می گردد. نظر به گسترش سیستم پایگاه داده و ایجاد ابزارهای متعدد برای ذخیره حجم بالایی از داده‌ها، داده کاوی به عنوان یکی از شاخه های علوم بین رشته ای توسعه روزافزونی یافته است. داده کاوی به جستجو و کشف مدل های گوناگون، چکیده سازی و اخذ مقادیر از مجموعه ای از داده های معلوم اطلاق می گردد

[Downloaded from jwmr.sanru.ac.ir on 2026-10-10]

(۹). دسته بندی و پیش بینی دو نوع عملیات داده‌کاوی برای تحلیل داده‌ها و استخراج مدل به منظور توصیف دسته های مهم داده ها، فهم و پیش بینی رفتار آن ها در آینده است. مدل های دسته ای در تحلیل داده‌های گسسته و طبقه ای به کار رفته و مدل های پیش بینی یا رگرسیون بیشتر روی داده های پیوسته کار می کنند (۱۱). تحقیقات متعددی روی کارایی روش های داده‌کاوی در مطالعات مربوط به منابع آب صورت گرفته، که به تعدادی از آنها اشاره می شود.

میشرا و دسای (۱۲) با استفاده از مدل های آماری به پیش بینی خشکسالی در حوزه کانساباتی هند پرداختند و سپس در سال ۲۰۰۶ کارایی روش های آماری و شبکه های عصبی مصنوعی در پیش بینی خشکسالی را با هم مقایسه نمودند که نتایج نشان دهنده برتری شبکه های عصبی مصنوعی نسبت به روش های آماری بود. در تحقیق دیگری که در همین سال توسط نامبردگان انجام گرفت پیش بینی خشکسالی با استفاده از سه روش آماری، شبکه عصبی مصنوعی و مدل های درآمیخته (هیبریدی) با هم مقایسه شد که مدل های هیبرید نتایج برتری نسبت به دو روش دیگر داشت. مرید و همکاران (۱۳) با به کارگیری شبکه های عصبی و استفاده از برخی شاخص های خشکسالی به پیش بینی خشکسالی در استان تهران پرداختند که در نهایت نمایه خشکسالی مؤثر در پیش بینی خشکسالی نسبت به سایر روش ها مؤثرتر واقع شد.

فاتحی مرج و همکاران (۶) با به کارگیری مدل شبکه عصبی مصنوعی و استفاده از شاخص های SOI و ۴، ۳ NINO به پیش بینی خشکسالی پاییزه یک فصل زودتر از وقوع پرداختند. آنها داده های نام برده را از ایستگاه ارومیه واقع در حوزه دریاچه ارومیه مورد استفاده قرار دادند و سپس نتایج حاصل از آن را با مدل رگرسیون چند متغیره، مقایسه کردند. آنها بدینوسیله موفق به بررسی روند تغییرات بارش شدند (۶).

سلسو آگستو و همکاران (۴) پیش بینی خشکسالی را با استفاده از شبکه های عصبی مصنوعی پیش تغذیه برای سه ناحیه همگن بارش درودخانه سانفرانسسکو برزیل انجام داد. در این مطالعه از داده‌های ۲۴۸ ایستگاه بارندگی استفاده شد. نتایج بدست آمده نشان داد که شبکه های عصبی مصنوعی روشی مناسب برای انواع پیش بینی هاست و زمانی که از شاخص بارش استاندارد جهت پیش بینی ها انجام می شود نتایج بهبود بخشیده می شود و در نهایت نتایج نهایی تحت تأثیر زون های مختلف هیدرولوژیکی نمی باشد.

افخمی و همکاران (۲) در تحقیقی با عنوان «بررسی تأثیر عناصر اقلیمی بر افزایش دقت روش شبکه عصبی مصنوعی در پیش بینی خشکسالی منطقه یزد» نقش تعداد و نوع

[Downloaded from jwmr.sanru.ac.ir on 2026-10-10]

عوامل اقلیمی را در دقت پیش بینی خشکسالی با شبکه های عصبی مصنوعی مورد ارزیابی قرار دادند. حبیبی پور (۷) با استفاده از آمار ایستگاه سینوپتیک یزد (در دوره آماری ۱۳۸۴- ۱۳۳۲) به بررسی کارایی درخت تصمیم در پیش بینی خشکسالی پرداخت. سپس برای بررسی کارایی الگوریتم درخت تصمیم در پیش بینی خشکسالی، مدل در چهار حالت استفاده از داده های خام، استفاده از داده های نرمال، استفاده از میانگین متحرک سه ساله و استفاده از میانگین متحرک پنج ساله اجراگردید. نتایج این تحقیق نشان داد که درخت تصمیم رگرسیونی با دقت قابل قبولی قادر به پیش بینی میانگین متحرک بارش (سه ساله و پنج ساله) می باشد (۷).

جی لون چیانگ و چیانگ وی (۵) پیش بینی خشکسالی مخزن را با استفاده از شبکه های عصبی مصنوعی انجام داد. در این مطالعه متغیر های مربوط به خشکسالی مخزن از جمله مخزن ذخیره و جریانات زیرقشری استفاده شد. جهت آنالیز و مقایسه نتایج خشکسالی مخزن از احتمال ماکزیمم و روش بیضی به کار گرفته شد. نتایج خروجی، عملکرد بهتر شبکه عصبی را نسبت به روش های دیگر نشان داد. به طوری که استفاده از شبکه عصبی مصنوعی اطلاعات مفیدی را جهت مدیریت و اداره مخازن فراهم می آورد. جهت انجام این تحقیق ۷ مدل از شبکه های عصبی شامل MLP[1]، GFF[2]، MNN[3]، PCA[4]، RN۴[5]، TLRN[6] مورد استفاده قرار گرفت در تمام مدل ها بارش نرمال ماهیانه به عنوان ورودی به شبکه معرفی شد و بارش ماهیانه یک سال آینده پیش بینی شد. لازم به ذکر است در این تحقیق از الگوریتم آموزشی موسوم به قانون یادگیری ممنتم استفاده شد. بعد از بررسی مدل های نامبرده، دو شبکه TLRN و RN مورد استفاده قرار گرفت که در نهایت شبکه TLRN بهترین کارایی را در پیش بینی خشکسالی از خود نشان داد. بهترین تابع فعال تابع تانژانت هایپربولیک تشخیص داده شد. از بین عناصر دینامیک در شبکه ها نوع لاگور (Laguarre) و تأخیر زمانی(TDNN) سازگاری بهتری را از خود نشان دادند.

لازم به ذکر است هدف از این تحقیق بررسی کارایی دو روش داده کاوی (شبکه عصبی مصنوعی و درخت تصمیم) در پیش بینی بارش در حالت استفاده از داده‌های اصلی و مقایسه آن با کارکرد مدل در شرایط استفاده از میانگین متحرک سه ساله می باشد. در این حالت میتوان به بررسی تأثیر عملیات پردازشی بر کارایی مدل درخت تصمیم و شبکه عصبی مصنوعی در پیش بینی میزان بارش و در نهایت رخداد خشکسالی پرداخت.

مواد و روش ها

معرفی روش مورد استفاده

در این تحقیق از دو روش شبکه عصبی مصنوعی و درخت تصمیم رگرسیونی استفاده شد، که در زیر به تشریح کلی آنها پرداخته می شود:

1- Multi Layer Perspetron
2- Generalized Feed Forward
3- Modular Neural Network
4- Principal Component Analysis
5- Recurrent Network
6- Time Lag Recurrent Network

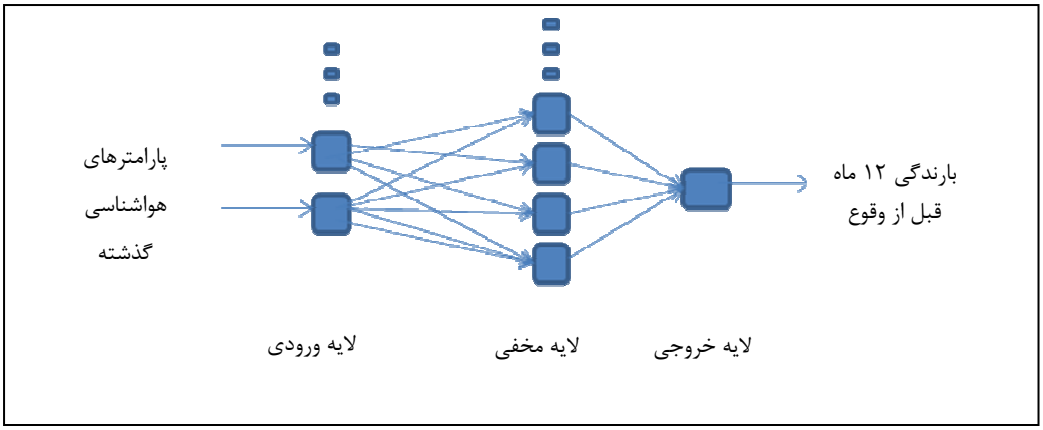
[Downloaded from jwmr.sanru.ac.ir on 2026-10-10]

شبکه عصبی مصنوعی

ایده شبکه عصبی مصنوعی را می توان به صورت ذیل تشریح نمود:

- مجموعه ای از گره ها (واحدها یا نرون ها یا عناصر محاسباتی).
- هر گره ورودی و خروجی دارد.
- هر گره براساس تابعی خاص محاسبه ساده ای انجام می دهد.
- بین گره ها، اتصالات موزون وجود دارد.
- اتصالات براساس معماری شبکه مشخص می شوند.
- نتیجه یک شبکه تابعی بسیار پیچیده از ارتباطات موزون می باشد.

روش شبکه عصبی یک روش نوین در حل مسائل پیچیده ای می باشد که یا الگوریتم مشخص برای حل آنها وجود ندارد و یا از روش های متعارف دارای راه حلی بسیار طولانی و زمان بر هستند. این روش در سال های اخیر در مسائل مختلف منابع آب مورد استفاده قرار گرفته است و بیشترین کاربرد آن در مورد مسائلی بوده است که یک ارتباط غیرخطی میان پارامترهای معلوم و مجهول آن مسائل وجود دارد، زیرا یکی از قابلیت های مهم شبکه عصبی مصنوعی، درک رفتار غیرخطی یک سیستم می باشد (۸). شکل ۱ مدل مفهومی شبکه های عصبی مصنوعی را نشان می دهد.



شکل ۱- مدل مفهومی شبکه های عصبی مصنوعی.

ساختارهای مختلفی برای شبکه عصبی مصنوعی وجود دارد، که در این تحقیق دو ساختار از آنها شامل شبکه های عصبی مصنوعی برگشتی و شبکه های عصبی مصنوعی برگشتی با تأخیر زمانی مورد استفاده قرار گرفته است. این شبکه ها از نوع دینامیک هستند که دارای لایه‌هایی با اتصال برگشتی محلی می‌باشند. از ویژگی‌های مهم این شبکه‌ها وجود یک حافظه[1] بی‌نهایت عمیق است که بین لایه ورودی و میانی یا بین لایه میانی و خروجی و یا بین عوامل موجود در لایه میانی قرار گرفته است و قادر است روابط بین ورودی و خروجی‌ها را به صورت لحظه‌ای و آنی تخمین بزند. میزان نرون‌های لایه مخفی متناسب با ۲n

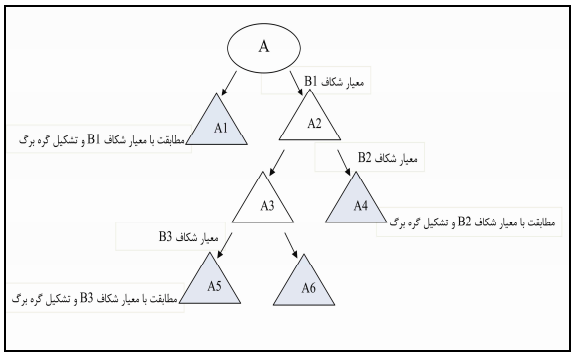
1- Unit Memory

[Downloaded from jwmr.sanru.ac.ir on 2026-10-10]

طراحی شد که n متناسب با تعداد عوامل ورودی به مدل می‌باشد. همچنین از توابع تانژانت هایپربولیک و سیگموئید به عنوان توابع فعال در لایه میانی و خروجی استفاده شد.

درخت تصمیم[1]

درخت تصمیم‌گیری یکی ازروشهای داده‌کاوی و از ابزارهای قوی و متداول برای دسته‌بندی و پیش‌بینی می‌باشد که برخلاف شبکه‌های عصبی به تولید قانون می‌پردازد. یعنی درخت تصمیم‌گیری پیش‌بینی خود را در قالب یکسری قوانین توضیح می‌دهد (شکل ۲). در حالی که در شبکه‌های عصبی تنها پیش‌بینی بیان می‌شود و چگونگی آن در خود شبکه پنهان باقی می‌ماند. علاوه بر آن در درخت تصمیم‌گیری بر خلاف شبکه‌های عصبی، می‌توان از داده‌های غیرعددی نیز استفاده نمود.



شکل ۲- نمونه‌ای از یک درخت تصمیم ساده.

ساختار این درخت بر سه اصل استوار است:

۱- تعیین مجموعه‌ای از پرسش ها به شکل $x \le d$? که در آن x یک متغیر مستقل و d یک مقدار ثابت است و جواب هر پرسش آری/ نه است.

۲- بهترین معیار شکاف برای انتخاب بهترین متغیرمستقل برای شکاف.

۳- ایجاد آمار خلاصه برای گره پایانی (۳).

معیار مورد استفاده در درختان تصمیم رگرسیونی، کمترین انحراف مربعات (LSD)[2] نام دارد. این معیار براساس کمیت خطا که به صورت زیر محاسبه می شود، مورد ارزیابی قرار می گیرد:

رابطه (۱) $$SS(t) = \sum_{i=1}^{Nt} (y_i(t) - \overline{y}(t))^2$$

$SS(t)$: مجموع مربعات خطا

N_t: تعداد سنجه ها (داده‌ها) در گره برگ t.

$y_i(t)$: مقدار خروجی (متغیر هدف در هر گره برگ).

$\overline{y}(t)$: میانگین مقادیر متغیر هدف برای همه گره‌ها.

حال متغیر ورودی S زمانی بهترین متغیر برای ایجاد شاخه در گره t می‌باشد که مقدار Q(s,t) را بیشینه نماید.

رابطه (۲) $$Q(s,t) = SS(t) - SS(t_R) - SS(t_L)$$

که در آن $SS(t_R)$ و $SS(t_L)$ به ترتیب میزان SS(t) در شاخه سمت راست و سمت چپ گره t می‌باشد.

آماده سازی داده‌ها

1- Decision Tree

2- Least Square Deviation

[Downloaded from jwmr.sanru.ac.ir on 2026-10-10]

در این تحقیق از آمار ایستگاه همدیدی یزد برای بازه زمانی ۱۳۸۴-۱۳۳۲ استفاده شد. این ایستگاه در شهرستان یزد واقع است و از سال ۱۳۴۰ شروع به کار نموده و در ارتفاع ۱۲۳۰متری از سطح دریا و طول جغرافیایی ′۱۷ °۵۴ و عرض جغرافیایی ′۵۳ °۳۱ قرار دارد. این ایستگاه با بالغ بر ۵۰ سال آمار، دارای بلندترین طول دوره آماری در بین ایستگاه‌های استان یزد می‌باشد.

در ابتدا آمار و اطلاعات مربوط به ایستگاه اخذ و از طریق آزمون توالی، همگنی داده‌ها بررسی شد. سپس ۱۲ ایستگاه هواشناسی استان که ضریب همبستگی پارامتر بارش آن ها با ایستگاه یزد در سطح ۱ درصد معنی دار بود شناسایی گردید. این ایستگاهها عبارتند از: ابرکوه، اردکان، خرانق، مزرعه‌نو عقدا، ساغند، حاج‌آباد زرّین، قطروم، باجگان، پاچنار تفت، نصرآباد، یزد و خویدک. با استفاده از روش نسبت نرمال و از طریق آمار و اطلاعات این ایستگاه ها، بازسازی کمبودهای آماری ایستگاه یزد انجام گردید. پس از آن، داده‌ها به نسبت ۷۰ به ۳۰ به دو دسته آموزش و آزمون تقسیم شدند. انتخاب داده‌های آموزش و آزمون به صورت سیستماتیک انجام شد ضمن اینکه هیچ یک از داده‌های آزمون در مرحله آموزش مورد استفاده قرار نگرفتند.

در دو مدل درخت تصمیم و شبکه عصبی مصنوعی، از پارامترهای بارش، میانگین دما، دمای بیشینه، رطوبت نسبی، سرعت باد و جهت باد به عنوان ورودی مدل استفاده و شبیه‌سازی‌ها برای پیش‌بینی بارش ۱۲ ماه قبل از وقوع صورت گرفت. متغیرهای مستقل مورد استفاده در این تحقیق و علائم اختصاری هر یک از این متغیرها در جدول ۱ نمایش داده شده است.

جدول ۱- معرفی پارامترهای ورودی مدل (ایستگاه یزد- دوره آماری ۱۳۳۲ لغایت ۱۳۸۴)

میانگین	حداقل	حداکثر	پارامتر ورودی	علامت اختصاری
۴/۴۷	۰	۷۰	بارش (میلیمتر)	P
۲۰/۱۸	-۰/۸	۳۵/۵	دمای متوسط (درجه سانتیگراد)	T_{mean}
۲۷/۲۵	۵/۸	۴۲/۶	دمای ماکزیمم (درجه سانتیگراد)	T_{max}
۲۲/۸۴	۸	۷۴	رطوبت نسبی (درصد)	H
۵/۷	۱/۶	۱۱/۱	سرعت شدید ترین باد (متر بر ثانیه)	W_s
۲۵۶/۶	۱۰	۳۶۰	جهت باد غالب (درجه)	W_d

اجرای مدل

در مجموع مدل شبکه عصبی مصنوعی و مدل درخت تصمیم در دو حالت کلی اجرا شد. در حالت اول از اصل داده‌ها استفاده شد، به عبارت دیگر روی داده‌ها هیچ گونه عملیات پردازشی انجام نشد و ورودی‌ها و خروجی همان داده‌های خام بودند که به مدل معرفی شدند. در حالت دوم از میانگین لغزان سه ساله داده‌ها برای اجرای مدل استفاده گردید. استان یزد یک استان خشک می‌باشد که از

[Downloaded from jwmr.sanru.ac.ir on 2026-10-10]

خصوصیات بارز آن نوسانات شدید بارش از یک سال به سال دیگر می‌باشد به گونه‌ای که داده‌های بارش از هیچ گونه روند منظمی برخوردار نمی‌باشند. از این رو پس از انجام چندین شبیه‌سازی برای بالا بردن ضریب کارایی در پیش‌بینی بارش، به جای اصل داده‌های ماهیانه از میانگین متحرک سه ساله آنها استفاده شد. به این ترتیب میانگین متحرک سه ساله برای کل پارامترهای ورودی مدل محاسبه شد و میانگین متحرک سه ساله بارش برای یک سال بعد پیش‌بینی گردید.

ارزیابی کارایی مدل

برای ارزیابی کارایی مدل های استفاده شده در این تحقیق از دو پارامتر آماری به شرح ذیل استفاده گردید:

رابطه (۳)
$$R = \sqrt{1 - \frac{\sum_{i=1}^{n}(Oi - Pi)}{\sum_{i=1}^{n}(Oi - \overline{O})}}$$

رابطه (۴)
$$RMSE = \left(\frac{\sum_{i=1}^{n}(p_i - o_i)^2}{n} \right)^{1/2}$$

در این روابط O و P به ترتیب مقادیر مشاهده شده و شبیه‌سازی شده، $\overline{O}$ و $\overline{P}$ میانگین مقادیر مشاهده شده و شبیه‌سازی شده و n تعداد نمونه‌ها می‌باشد.

نتایج و بحث

نتایج مدل درخت تصمیم در شرایط استفاده از اصل داد ه ها

در این حالت، مدل با ۶ متغیر مستقل شامل بارش، میانگین دما، دمای بیشینه، رطوبت نسبی، سرعت باد و جهت باد و ۶۲۰ سطر داده اجرا شد. لازم به ذکر است که از این تعداد ۴۳۴ سطر داده در مرحله آموزش و ۱۸۶ سطر داده در مرحله آزمون استفاده گردید. قوانین ایجاد شده در این حالت عبارتند از:

IF $T_{max} \leq 26.2$ then P- lead12=9.38
IF $T_{max} > 26.2$ and $T_{max} \leq 33.2$ and $W_d \leq 245$ then p-lead12=10.06
IF $T_{max} > 26.2$ and $T_{max} \leq 33.2$ and $W_d > 245$ and $T_{mean} \leq 24.05$ and $p \leq 1.8$ then p-lead12=0.523
IF $T_{max} > 26.2$ and $T_{max} \leq 33.2$ and $W_d > 245$ and $T_{mean} \leq 24.05$ and $p > 1.8$ then p-lead12=3.29
IF $T_{max} > 26.2$ and $T_{max} \leq 33.2$ and $W_d > 245$ and $T_{mean} > 24.05$ then p-lead12=4.59
IF $T_{max} > 26.2$ and T_{max} x>33.2 then p-lead12=0.27

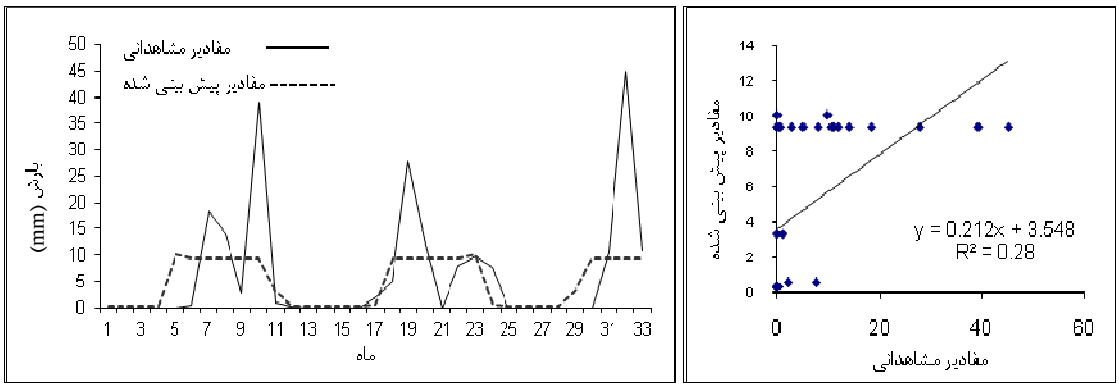
قوانین ایجاد شده توسط مدل درخت تصمیم نشان می دهد که مدل درخت تصمیم در شرایط استفاده از اصل داده‌ها، از بین شش پارامتری که در اختیار داشته، تنها از پارامتر بارش، میانگین دما، دمای بیشینه و جهت باد برای پیش بینی استفاده کرده است. مقادیر خطا برای مدل درختی در این حالت محاسبه و نتایج آن در جدول ۲ ارائه شده است.

جدول۲- نتایج مدل درخت تصمیم رگرسیونی در حالت استفاده از اصل داده ها

ردیف	متغیر مستقل	متغیر هدف	R	RMSE
حالت اول	$P\text{-}T_{mean}\text{-}T_{max}\text{-}W_d$	P_1	۰/۵۳	۸/۲

[Downloaded from jwmr.sanru.ac.ir on 2026-10-10]

نمودار نقطه ای و نمودار خطی مربوط به نتایج مدل درخت تصمیم و مقادیر واقعی مربوط به حالت استفاده از اصل داده در شکل ۳ نمایش داده شده است.



شکل ۳- نمودار تغییرات مقادیر پیش بینی شده بارش در مقابل مقادیر واقعی در شرایط استفاده از اصل داده ها- مدل درخت تصمیم رگرسیونی.

هر چند مدل درخت تصمیم از نظر آماری در حالت استفاده از اصل داده‌ها معنی دار بوده، ولی مشاهده می شود که مدل در این حالت در غالب دوره های زمانی۱۲ ماهه پیش بینی های خود را در قالب دو مقدار عددی حداقل و حداکثر انجام داده است. به عبارتی درخت ایجاد شده با استفاده از اصل داده‌ها روند پیش بینی نسبتاً ثابتی داشته و نتوانسته تغییرات را به خوبی شبیه سازی نماید.

نتایج مدل درخت تصمیم در حالت استفاده از میانگین متحرک سه ساله داد ه ها

در این حالت برای اجرای مدل، ۶ متغیر مستقل شامل میانگین متحرک سه ساله داده های بارش، دمای متوسط، دمای بیشینه، رطوبت نسبی، سرعت باد و جهت باد با ۵۹۶ سطر داده در اختیار سیستم قرار گرفت که از این تعداد ۴۱۷ سطر داده در مرحله آموزش و ۱۷۹ سطر در مرحله آزمون استفاده گردید. قوانین درخت ایجاد شده در این حالت عبارتست از:

IF $p \leq 8.15$ and $p \leq 3.9$ and $T_{max} \leq 32.6$ and then P- lead12=2.6

IF $p \leq 8.15$ and $p \leq 3.9$ and $T_{max} > 32.6$ and $p \leq 1.3$ and $p \leq 0.3$ then P- lead12=0.06

IF $p \leq 8.15$ and $p \leq 3.9$ and $T_{max} > 32.6$ and $p \leq 1.3$ and $p > 0.3$ then P- lead12=0.06

IF $p \leq 8.15$ and $p \leq 3.9$ and $T_{max} > 32.6$ and $p > 1.3$ then P- lead12=1.6

IF $p \leq 8.15$ and $p > 3.9$ and $W_d \leq 261.7$ then P- lead12=7.4

IF $p \leq 8.15$ and $p > 3.9$ and $W_d > 261.7$ then P- lead12=5.1

IF $p > 8.15$ and $p \leq 2.19$ and $p \leq 15.07$ then P- lead12=10.5

IF $p > 8.15$ and $p \leq 2.19$ and $p > 15.07$ then P- lead12=15.4

IF $p > 8.15$ and $p > 2.19$ then P- lead12=23.9

در این حالت درخت سه پارامتر بارش، دمای ماکزیمم و جهت باد را در شبیه سازی

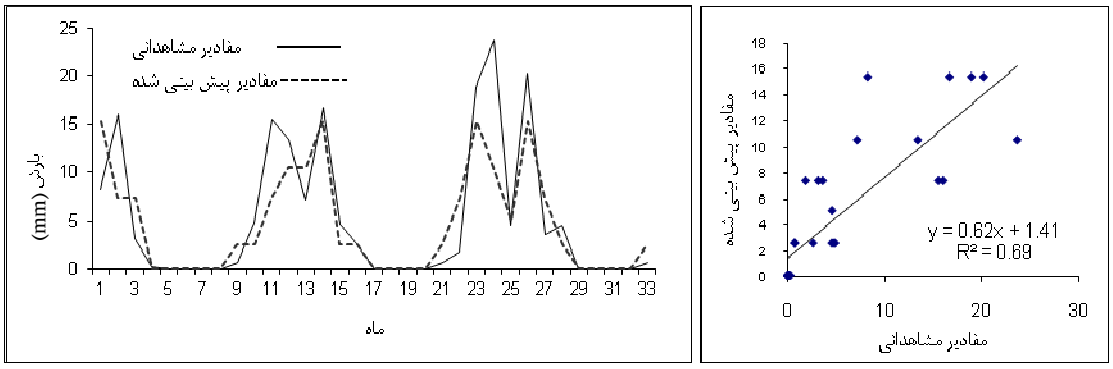
[Downloaded from jwmr.sanru.ac.ir on 2026-10-10]

لحاظ نموده است. علاوه بر آن قوانین ایجاد شده توسط درخت در حالت اول (استفاده از داده های اصلی) با حالت دوم (استفاده از میانگین متحرک سه ساله داده ها) متفاوت می باشد، که این موضوع حاکی از تأثیر تعدیل دامنه تغییرات داده های ورودی بر کارایی مدل درخت تصمیم گیری می باشد. کارایی درخت تصمیم ایجاد شده در این حالت نیز با استفاده از دو پارامتر آماری RMSE، R بررسی شد که مقادیر مربوط به این پارامتر ها در جدول ۳ درج شده است.

جدول ۳- نتایج مدل درخت تصمیم رگرسیونی با استفاده از میانگین متحرک سه ساله داده ها

ردیف	متغیر مستقل (میانگین متحرک سه ساله)	متغیر هدف (میانگین متحرک سه ساله)	R	RMSE
حالت دوم	$P- T_{max} -W_d$	P_3	۰/۸۳	۳/۴۴

نمودار نقطه ای و نمودار خطی مربوط به نتایج مدل درختان تصمیم و مقادیر واقعی مربوط به حالت استفاده از میانگین متحرک سه ساله داده‌ها در شکل ۴ نمایش داده شده است.



شکل۴- نمودار تغییرات مقادیر پیش بینی شده بارش در مقابل مقادیر واقعی با استفاده از میانگین متحرک سه ساله- مدل درخت تصمیم رگرسیونی.

بررسی ها نشان داد مدل درخت تصمیم در حالت استفاده از میانگین لغزان داده‌ها هم معنی دار می باشد. علاوه بر آن توانایی مدل در پیش بینی خشکسالی به میزان قابل توجهی افزایش یافته است (جداول ۲ و ۳).

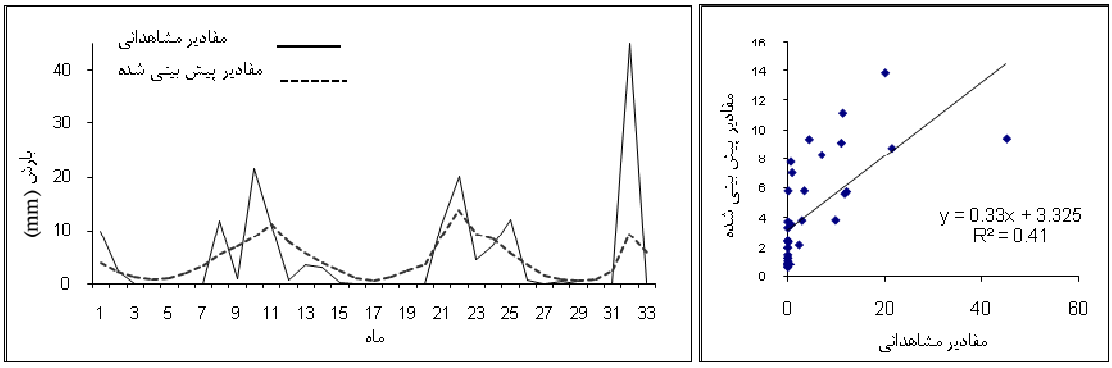
نتایج مدل شبکه عصبی مصنوعی در شرایط استفاده از اصل سنجه ها (داده ها)

در این تحقیق از دو ساختار شبکه عصبی شامل شبکه های عصبی مصنوعی برگشتی و شبکه های عصبی مصنوعی برگشتی با تأخیر زمانی استفاده شد. نتایج بهترین شبیه سازی های صورت گرفته با استفاده از شبکه عصبی مصنوعی در حالت استفاده از اصل داده‌ها در جدول ۴ آمده است.

[Downloaded from jwmr.sanru.ac.ir on 2026-10-10]

جدول ۴- نتایج مدل شبکه عصبی مصنوعی در حالت استفاده از اصل داده ها

ردیف	نوع مدل	ترکیب ورودی	تابع لایه میانی	تابع لایه خروجی	Epoch	نوع خروجی (بارش ۱۲ ماه قبل از وقوع)	R	RMSE
۱	RN	$P\text{-}T_{mean}$	sigxon	tanhaxon	۲۵۰۰۰	P_1	۰/۴۲	۸/۰۱
۲	RN	$P\text{-}T_{mean}\text{-}H$	sigxon	Tanhaxon	۲۵۰۰۰	P_1	۰/۳۶	۱۱/۹۶
۳	RN	$P\text{-}T_{mean}\text{-}W_s\text{-}W_d$	sigxon	Tanhaxon	۱۰۰۰۰	P_1	۰/۴۲	۸/۷۷
۴	RN	$P\text{-}H\text{-}W_s\text{-}W_d$	sigxon	Tanhaxon	۱۰۰۰۰	P_1	۰/۴۴	۸/۷۷
۵	RN	$P\text{-}T_{mean}\text{-}H\text{-}W_s\text{-}W_d$	sigxon	Tanhaxon	۱۰۰۰۰	P_1	۰/۴۰	۸/۱۱
۶	TLRN	$P\text{-}T_{mean}$	tanhaxon	Tanhaxon	۱۰۰۰۰	P_1	۰/۶۴	۵/۴
۷	TLRN	$P\text{-}T_{mean\text{-}H}$	tanhaxon	Tanhaxon	۱۰۰۰۰	P_1	۰/۴۵	۶/۳
۸	TLRN	$P\text{-}T_{mean}\text{-}W_s\text{-}W_d$	tanhaxon	Tanhaxon	۱۰۰۰۰	P_1	۰/۵۸	۷/۵
۹	TLRN	$P\text{-}H\text{-}W_s\text{-}W_d$	tanhaxon	Tanhaxon	۱۰۰۰۰	P_1	۰/۴۵	۷/۰۲
۱۰	TLRN	$P\text{-}T_{mean}\text{-}H\text{-}W_s\text{-}W_d$	tanhaxon	tanhaxon	۱۰۰۰۰	P_1	۰/۵۱	۶/۴



شکل ۵- نمودار تغییرات پیش بینی شده بارش در مقابل مقادیر واقعی در شرایط استفاده از اصل داده ها، ترکیب بارش- دمای متوسط (مدل شبکه عصبی مصنوعی).

بررسی ها نشان داد که از نظر آماری مدل شبکه عصبی مصنوعی هم در شرایط استفاده از اصل داده‌ها معنی دار می باشد و بهترین نتیجه در شرایط استفاده از میانگین دما و بارش به عنوان ورودی مدل به دست آمد.

نتایج مدل شبکه عصبی مصنوعی در شرایط استفاده از میانگین متحرک سه ساله داد ه ها

نتایج بهترین شبیه سازی های صورت گرفته در این حالت در جدول ۵ آمده است. بهترین نتیجه در شرایطی حاصل شده که از بارش و دمای بیشینه به عنوان ورودی مدل استفاده شد.

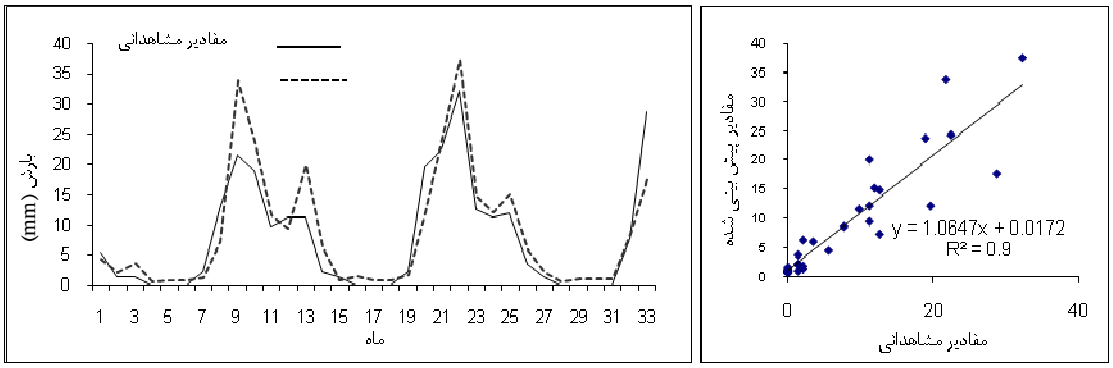
معنی داری مدل در این حالت از نظر آماری بررسی و نتایج حاکی از معنی دار بودن مدل می باشد. با توجه به نمودارهای ترسیم شده (اشکال ۳، ۴، ۵ و ۶)، می توان گفت در شبیه سازی های صورت گرفته، مدل در پیش بینی مقادیر کمینه از توانایی بیشتری برخوردار است. لذا با توجه به اینکه هدف، ارزیابی هر چه بهتر وضعیت خشکسالی در آینده از طریق مؤلفه بارندگی است، توانایی بیشتر دو مدل مورد استفاده در پیش بینی مقادیر کم بارش به نسبت مقادیر زیاد آن قابل تأمل، مورد نظر و مطلوب می باشد. بهترین

[Downloaded from jwmr.sanru.ac.ir on 2026-10-10]

نتـایج حاصـل از دو مـدل مـورد اسـتفاده در جدول ۶ آمده است.

جدول ۵- نتایج مدل شبکه عصبی مصنوعی با استفاده از میانگین متحرک سه ساله داده ها

ردیف	نوع مدل	ترکیب ورودی (میانگین متحرک سه ساله)	تابع لایه میانی	تابع لایه خروجی	Epoch	نوع خروجی (میانگین متحرک سه ساله)	R	RMSE
۱	TLRN	$P\text{-}T_{mean}\text{-}H\text{-}W_s\text{-}W_d$	tanhaxon	tanhaxon	۲۳۰۰	P_3	۰/۸۴	۴/۹
۲	TLRN	P-H	tanhaxon	tanhaxon	۱۰۰۰۰	P_3	۰/۸۶	۵/۶
۳	TLRN	$P\text{-}T_{mean}\text{-}H\text{-}W_s$	tanhaxon	tanhaxon	۷۰۰۰	P_3	۰/۸۷	۵/۶
۴	TLRN	$P\text{-}H\text{-}W_s\text{-}W_d$	tanhaxon	tanhaxon	۳۰۰۰	P_3	۰/۸۸	۴/۹
۵	TLRN	$P\text{-}T_{max}\text{-}H\text{-}W_s$	tanhaxon	tanhaxon	۸۶۰	P_3	۰/۸۸	۵/۶
۶	TLRN	$P\text{-}T_{mean}\text{-}W_s\text{-}W_d$	tanhaxon	tanhaxon	۳۰۰۰	P_3	۰/۹	۴/۹
۷	TLRN	$P\text{-}T_{mean}\text{-}H$	tanhaxon	tanhaxon	۲۳۰۰	P_3	۰/۹	۴/۹
۸	TLRN	$P\text{-}W_s$	tanhaxon	tanhaxon	۱۰۰۰۰	P_3	۰/۹۲	۴/۲
۹	TLRN	$P\text{-}T_{mean}\text{-}H$	tanhaxon	tanhaxon	۱۰۰۰۰	P_3	۰/۹۳	۴/۲
۱۰	TLRN	$P\text{-}T_{mean}$	tanhaxon	tanhaxon	۱۸۰۰۰	P_3	۰/۹۵	۴/۹
۱۱	TLRN	$P\text{-}T_{max}$	tanhaxon	tanhaxon	۲۲۰۰۰	P_3	۰/۹۵	۳/۵



شکل ۶- نمودار تغییرات پیش بینی شده بارش در مقابل مقادیر واقعی با استفاده از میانگین متحرک سه ساله، ترکیب بارش- دمای بیشینه (مدل شبکه عصبی مصنوعی).

جدول ۶- بهترین نتایج حاصل از دو مدل درخت تصمیم و شبکه عصبی مصنوعی

مدل مورد استفاده	متغیر ورودی (میانگین متحرک سه ساله)	متغیر هدف (میانگین متحرک سه ساله)	میزان تغییر نتایج نسبت به حالت استفاده از اصل داده ها (درصد)	
			R	RMSE
بهترین نتیجه در مدل درخت تصمیم	$P\text{-}T_{max}\text{-}W_d$	P	۷۳/۹	-۱۳۸/۴
بهترین نتیجه در مدل شبکه عصبی مصنوعی	$P\text{-}T_{max}$	P	۵۵/۵	-۵۴/۳

در مجموع در ایستگاه یزد، دو مدل داده کاوی درختان تصمیم و شبکه عصبی مصنوعی در پیش‌بینی خشکسالی در شرایط استفاده از اصل داده‌ها ناموفق بود. به نظر می‌رسد دامنه تغییرات این داده‌ها برای استفاده در مدل های داده‌کاوی زیاد می‌باشد و همین موضوع منجر به کاهش کارایی مدل در پیش‌بینی می‌گردد. بهتر شدن کارایی هر دو

[Downloaded from jwmr.sanru.ac.ir on 2026-10-10]

مدل در پیش بینی به هنگام استفاده از داده های میانگین متحرک مؤید این موضوع می باشد. به عبارتی هر دو مدل درخت تصمیم و شبکه عصبی مصنوعی در شرایط استفاده از میانگین متحرک داده ها نتایج قابل قبولی داشته اند. ضمن اینکه اعمال فرآیند میانگین متحرک روی داده های ورودی به نحو چشمگیری بر کارایی هر دو مدل و به ویژه مدل درخت تصمیم رگرسیونی تأثیرگذار بوده است.

لازم به ذکر است در بهترین شبیه سازی صورت گرفته توسط هر دو مدل، پارامترهای بارش و دمای ماکزیمم دخیل بوده است که این موضوع نشان دهنده اهمیت این پارامترها در پیش بینی فرآیند خشکسالی می باشد.

در چهار شبیه سازی صورت گرفته در این تحقیق، زمانی که از میانگین متحرّک سه ساله داده‌ها برای اجرای مدل شبکه عصبی مصنوعی استفاده گردید، ترکیب "بارش، دمای بیشینه" به عنوان ورودی های مدل، منجر به کسب دقیقترین نتایج شد (جدول ۷).

جدول ۷- بهترین نتایج حاصل از دو مدل درخت تصمیم و شبکه عصبی مصنوعی

مدل مورد استفاده	متغیر ورودی (میانگین متحرک سه ساله)	متغیر هدف (میانگین متحرک سه ساله)	R	RMSE
درخت تصمیم	P- T_{max} -W_d	P	۰/۸۳	۳/۴۴
شبکه عصبی مصنوعی	P- T_{max}	P	۰/۹۵	۳/۵

هرچند دقت مدل شبکه عصبی مصنوعی اندکی بیشتر از مدل درخت تصمیم می باشد. لیکن شبکه های عصبی مصنوعی به مثابه مدل های جعبه سیاه عمل کرده و نحوه عملکرد مدل در آن ها مشخص نیست، ولی مدل درخت تصمیم با تولید قوانین در زمره مدلهای جعبه سفید تلقی می گردد.

نتایج بدست آمده از این تحقیق با نتایج تحقیقات مشیرا و مرید (۱۲ و ۱۳) که به ترتیب خشکسالی را با استفاده از شبکه عصبی مصنوعی در حوزه کانساباتی هند و تهران انجام دادند مقایسه شد. میشرا پیش بینی خشکسالی را با استفاده از شاخص SPI ۳، ۶، ۱۲ و ۲۴ ماهه و در تأخیرهای زمانی ۱ تا ۶ ماهه انجام داده است. علیرغم اینکه منطقه تحقیق وی منطقه ای مرطوب با دامنه تغییرات کم داده‌ها بوده، بهترین نتیجه در شرایط استفاده از SPI ۲۴ ماهه و تأخیر زمانی یک ماهه (R^2 =۰/۹۲-۰/۹۳) به دست آمده است. با افزایش تأخیر زمانی میزان R^2 کاهش یافته، به طوری که میزان R^2 با تأخیر زمانی ۶ ماهه به ۰/۵۹-۰/۶۳ تقلیل یافته است.

مرید (۱۳) در منطقه تهران با استفاده از دو شاخص SPI و EDI و چندین مدل شبکه عصبی خشکسالی را برای ۱۲-۱ ماه آینده پیش بینی کرد. در تحقیق وی بهترین نتایج در تأخیر زمانی شش ماهه (R^2=۰/۶۶-۰/۹۸) حاصل شد که این ضریب با افزایش زمان تأخیر

[Downloaded from jwmr.sanru.ac.ir on 2026-10-10]

کاهش یافت و در تأخیر زمانی ۱۲ ماهه برای شاخص SPI به R^2=۰/۵۴-۰/۶۲ و برای شاخص EDI به R^2 =۰/۵۵-۰/۷۲ کاهش یافت.

لازم به ذکر است در رابطه با پیش بینی بارش و به تبع آن ارزیابی خشکسالی قبل از وقوع، تا کنون موفقیت قابل توجهی در سطح جهان حاصل نشده است. این مسئله هنگامی که فاصله زمانی بین پیش بینی و وقوع خشکسالی افزایش می یابد شدیدتر می شود. علاوه بر آن در مناطق خشک و نیمه خشک (که پراکنش داده های اقلیمی به ویژه بارش نسبت به میانگین دراز مدت آن زیاد می باشد) مدل های پیش بینی از ناتوانی بیشتری برخوردار می گردند. نتیجه این تحقیق نشان داد که پیش پردازش مناسب داده‌های مرتبط با خشکسالی همراه با به کارگیری مدل های مناسب داده‌کاوی می تواند افق جدیدی در فرایند پیش بینی پدیده خشکسالی ایجاد کرده و دقت پیش بینی ها را به مقدار قابل توجه ای ارتقاء بخشد. البته لازم است تحقیقات بیشتر در این زمینه در مناطق مختلف با شرایط اقلیمی و داده‌ای متفاوت صورت پذیرد، تا بتوان نسبت به انعطاف پذیری و سازگاری مدل های استفاده شده در شرایط مختلف اطمینان حاصل کرد. بدیهی است اصلی ترین عامل در مدیریت پدیده های خشکسالی و کاهش خسارات مربوطه، امکان پیش بینی مناسب آن می باشد و هرگونه پیشرفتی در این زمینه حائز اهمیت ویژه است.

مقادیر بارندگی جزء بی‌ثبات‌ترین متغیرهای آب و هوایی مخصوصاً در مناطق خشک محسوب می‌شود و به لحاظ تأثیرات مستقیم بر خشکی، شاخص خوبی برای مطالعه خشکسالی می‌باشد. به همین علت هم غالباً اصلی ترین عامل جهت محاسبه شاخص های خشکسالی (SPI، RDI و ...) میزان بارش است. لذا هر مدلی که بتواند میزان بارش را با دقت کافی قبل از وقوع آن پیش بینی نماید، می تواند جهت ارزیابی شدت خشکسالی در آینده مورد استفاده قرار گیرد. بدیهی است اطلاع از شدت و وضعیت خشکسالی می تواند تصمیم گیران و جوامع تحت تأثیر را جهت آمادگی و مقابله با خسارات آن به شکل قابل توجهی یاری نماید. علاوه بر آن پارامتر های هواشناسی و به ویژه پارامتر بارش از پایگاه داده مناسبی در سازمان های مسئول زیربط برخوردار می باشد، لذا می توان از طریق داده کاوی به پردازش، شبیه سازی و تجزیه و تحلیل داده‌های مذکور و کشف روابط حاکم بر داده‌های اقلیمی موثر بر رخداد خشکسالی پرداخت. این در حالی است که نوسانات شدید بارش به خصوص در مناطق خشک، توانایی مدل های داده‌کاوی را در یادگیری فرایند حاکم بر وقوع خشکسالی تحت الشعاع قرار می دهد. از این رو پردازش داده‌ها به منظور افزایش کارایی این مدل ها می تواند در تحقق اهداف تحقیق مثمر ثمر قرار گیرد. نتایج این تحقیق نشان داد که اعمال فرآیند میانگین متحرک به عنوان یکی از عملیات موثر در تعدیل دامنه تغییرات داده‌ها، به نحو چشمگیری بر افزایش کارایی هر دو مدل مورد استفاده و به ویژه مدل درخت تصمیم

[Downloaded from jwmr.sanru.ac.ir on 2026-10-10]

رگرسیونی تأثیرگذار بوده است. نتایج به دست آمده نشان می دهد که هر دو روش درخت تصمیم رگرسیونی و روش شبکه عصبی مصنوعی در ایستگاه یزد، مدل هایی کارا در پیش‌بینی خشکسالی با استفاده از داده‌های پردازش شده می‌باشند، که در صورت تعدیل دامنه تغییرات داده‌های ورودی قادرند با ضریب اطمینان بالایی میزان بارش را ۱۲ ماه قبل از وقوع برآورد نمایند. به طور مشخص، استفاده از میانگین متحرّک سه ساله نسبت به حالت استفاده از اصل داده‌ها منجر به افزایش چشمگیر کارایی مدل های درخت تصمیم رگرسیونی و شبکه عصبی مصنوعی می‌شود. ضمن اینکه اعمال فرآیند میانگین متحرک در مدل های هوشمند نظیر درخت تصمیم منجر به تغییر ترکیب پارامترهای ورودی و قوانین ایجاد شده توسط مدل را تغییر داد.

منابع:

1. Afkhami, H. 2008. Evaluation and prediction of drought in Yazd area using artificial neural network. Thesis of M.Sc. Yazd university. 236 pp.
2. Afkhami, H., M.T. Dastorani, H. Malekinezhad and M.H. Mobin. 2010. Evaluation of the effects of climatic factors on ANN based drought prediction in Yazd area. Journal of sciences and Technology of Agriculture and Natural Resources. 51: 157-169.
3. Breiman, L., J. Friedman, R. Olshen and C. Stone. 1984. Classification and Regression Trees, Monterey, CA:Wadsworth and Brooks/cole Advanced Books and Software. 105 pp.
4. Celso Augusto, G., S. Santon Bruno Morais and B.L. Gustavo. 2009. Drought forecast using an artificial neural network for three hydrological zones in san Francisco river basin. Brazil. New Approaches to Hydrological Prediction in Data-Sparse Regions. IAHS. 333: 302-312.
5. Chiang, J.L. and W. Chiang. 2011. Drought prediction using artificial neural network. Geophysical Research Abstracts. Vol. 13, EGU 2011-2224.
6. Fatehi, A., M.H. Davoudi and N. Arman. 2008. Evaluation of drought using SOI and NINO3, 4 Index and Artificial Neural Network. The 5th National Conference on Watershed Management Engineering. 173 pp.
7. Habibipour, A. 2010. Prediction of drought in synoptic station of Yazd using decision trees. Thesis of M.Sc. Yazd university. 130 pp.
8. Ghazanfari, M., S. Alizadeh and T. Pour babak. 2008. Data mining and Discovering Knowledge. Press center of Ferdowsi university. 335 pp.
9. Kantardzic, M. 1998. Data Mining. 208 pp.
10. Kardvani, P. 2001. Drought and The methods to prevent it. Tehran university Press 335 pp.
11. Meshgani, A. and A. Nazemi. 2009. An introduction to data mining. Mashhad university Press 185 pp.
12. Mishra, A.K. and V.R. Desai. 2006. Drought forecasting using feed-forward recursive neural network. Ecological Modelling 198: 127-138.

[Downloaded from jwmr.sanru.ac.ir on 2026-10-10]

13. Morid, S., V. Smakhtin and K. Bagherzadeh. 2007. Drought forecasting using artificial neural networks and time serried of drought indices. Int. Journal of Climatology. 27: 2103-2111.
14. Wilhite, D.A. 1993. The enigma of drought, chapter 1, In: D.A. Wilhite(ed), Drought Assessment Management, and planning: Theory and case studies, Kluwer Academic publishers, Boston, MA, pp: 3-17.

[Downloaded from jwmr.sanru.ac.ir on 2026-10-10]

Evaluation of the Effects of Data range Modification on Efficiency of Regression Decision Tree and Artificial Neural Networks for Drought Prediction

A. Habibipoor[1], M.T. Dastorani[2], M.R. Ekhtesasi[2] and H. Afkhami[3]

1- M.Sc. Faculty of Natural Resources, Yazd University, (Corresponding author: habibipoor_azam@yahoo.com)
2- Associate Professor, Faculty of Natural Resources, Yazd University
3- M.Sc. Faculty of Natural Resources, Yazd University

Abstract

One of the effects of climate system change is the occurrence and intensification of drought phenomenon. Prediction of drought condition can play an important role in mitigation of its effects as well as effective management of the available water during the drought periods. Different approaches have been presented for evaluation of drought. Analysis of precipitation data is the general method for drought evaluation, as acceptable prediction of precipitation before its occurrence, would be necessary and effective for analysis of drought. The purpose of this research is the evaluation of the effect of data processing on applicability of two data mining models on drought prediction in Yazd station. In addition, during the recent decades some new computer based models have been developed for drought prediction and in most of the cases they have presented quite satisfactory results. In this research, prediction of precipitation, which is the main component on drought occurrence, has been carried out in Yazd synoptic meteorological station. Therefore, two data mining methods including Regression Decision Tree (RDT) and Artificial Neural Networks (ANN) have been used, and simulations were carried out in two different conditions. In the first condition, the measured values of some meteorological variables were used as inputs and the amount of precipitation was predicted 12 months in advance. However, In the second condition, 3-year moving average of data were the inputs of the models for prediction of precipitation amount 12 months before its occurrence. Finally for evaluation of the model performance in different conditions, statistical criterion including R and RMSE were employed. Results indicated that using moving average of data as inputs of the models has considerably improved the performance of the models. Both RDT and ANN methods are able to predict the amount of precipitation in Yazd station 12 months before its occurrence.

Keywords: Prediction, Drought, Artificial Neural Networks, Regression Decision Tree, Data mining

[Downloaded from jwmr.sanru.ac.ir on 2026-10-10]