

## ارزیابی کارایی چندین روش داده کاوی برای پیش بینی تبخیر (مطالعه موردی: ایستگاه سینوپتیک یزد)

- ❖ حمیده افخمی\*؛ دکتری علوم و مهندسی آبخیزداری، دانشکده منابع طبیعی و کویرشناسی، دانشگاه یزد
- ❖ اعظم حبیبی پور؛ دکتری علوم و مهندسی آبخیزداری، دانشکده منابع طبیعی و کویرشناسی، دانشگاه یزد
- ❖ محمدرضا اختصاصی؛ استاد دانشکده منابع طبیعی و کویرشناسی، دانشگاه یزد

### چکیده

تبخیر یکی از پارامترهای اقلیمی مهم در مناطق خشک است و نقش مهمی را در مدیریت منابع آب بازی می کند، به همین جهت آگاهی از مقدار تبخیر و مدل سازی آن به عنوان یکی از متغیرهای مهم هیدرولوژیکی در تحقیقات کشاورزی و حفاظت آب و خاک از اهمیت زیادی برخوردار است. در دهه های اخیر روش های هوش مصنوعی در تخمین و پیش بینی پدیده های غیرخطی توانایی بالایی از خود نشان داده است. در این تحقیق از سه روش مهم داده کاوی شامل شبکه عصبی مصنوعی، شبکه های استنتاج فازی و درخت تصمیم رگرسیونی جهت پیش بینی تبخیر ماهانه در ایستگاه سینوپتیک یزد استفاده شد. برای این منظور از ۸ متغیر هواشناسی در مقیاس ماهانه (متوسط کمینه دما، متوسط بیشینه دما، میانگین دما، ساعات آفتابی، سرعت باد، جهت باد، میانگین رطوبت نسبی و تبخیر) به عنوان ورودی مدل استفاده گردید. نتایج به دست آمده نشان داد هر سه مدل نامبرده قادرند با استفاده از پارامترهای اقلیمی مذکور به پیش بینی مقدار تبخیر ماهانه ۱۲ ماه بعد از وقوع بردازند ولی در میان سه مدل مورد استفاده، شبکه عصبی مصنوعی با ضریب همبستگی برابر با  $r=0.97$ ،  $RMSE=5.1$ ،  $MAE=3.3$  و  $ME=-0.48$  بهترین کارایی را از خود نشان داد. همچنین نتایج نشان داد در پیش بینی تبخیر، تفاوت قابل ملاحظه ای در زمان استفاده از داده های خام و داده های نرمال شده وجود ندارد و پردازش داده ها تأثیر چندانی در بهبود نتایج مدل ها نخواهد داشت.

**کلید واژگان:** پیش بینی، تبخیر، شبکه عصبی مصنوعی، استنتاج فازی، درخت تصمیم گیری، یزد.

## ۱. مقدمه

تبخیر یکی از متغیرهای اقلیمی بسیار مهم به خصوص در مناطق خشک به شمار می‌رود. این پدیده ناشی از عدم توازن فشار بخار بین سطح آب و اتمسفر می‌باشد. در مناطقی که میزان بارش نیز اندک باشد میزان تبخیر در کاهش منابع ذخیره و پایین انداختن سطح تراز آب نقش مؤثری ایفا می‌کند. بنابراین با توجه به آنچه بیان شد تخمین میزان تلفات تبخیر اهمیت زیادی در بخش کنترل منابع آب در سطح مزارع و همچنین در مقیاس منطقه‌ای دارد.

امروزه فرمول‌های تجربی متعددی در زمینه برآورد میزان تبخیر ارائه شده است. همچنین روش‌های متعددی وجود دارد که به صورت مستقیم و غیر مستقیم میزان تبخیر را از سطوح آزاد آب تخمین می‌زنند. به دلیل ماهیت پیچیده و غیرخطی فرایند تبخیر استفاده از فرمول‌های تجربی در برآورد میزان تبخیر از صحت کافی برخوردار نمی‌باشد. بنابراین با توجه به اهمیت برآورد صحیح میزان تبخیر استفاده از روش‌های دقیق برای شبیه‌سازی تبخیر ضروری به نظر می‌رسد [۶، ۴۰]. در دهه‌های اخیر تکنیک‌های هوش مصنوعی از جمله شبکه عصبی مصنوعی<sup>۱</sup> و شبکه‌های استنتاج فازی<sup>۲</sup> و درخت تصمیم<sup>۳</sup> در تخمین و پیش‌بینی پدیده‌های غیرخطی توانایی و انعطاف‌پذیری بالایی را از خود نشان داده است به طوری که اخیراً پژوهشگران با به کار بردن شبکه‌های عصبی در زمینه تبخیر و تبخیر تعرق به نتایج با اهمیتی در این زمینه دست‌یافته‌اند. سادهر و همکاران تبخیر از سطح تشتک را با استفاده از شبکه‌های پرسپترون انجام دادند و نشان دادند در شرایط کمبود داده‌های هواشناسی می‌توان تنها با استفاده از داده‌های دما نتایج رضایت‌بخشی در تخمین میزان تبخیر به دست آورد [۳۴]. کومار و همکاران نیز نتایج استفاده از شبکه عصبی در پیش‌بینی تبخیر را مثبت ارزیابی کردند [۲۰]. تراجوکیچ و همکاران پیش‌بینی میزان تبخیر و تعرق پتانسیل ماهانه را با استفاده از شبکه‌های عصبی انجام

دادند. بر این اساس میزان تبخیر و تعرق نیز با استفاده از روش پنمن-مانتیس محاسبه گردید و سپس با استفاده از سایر پارامترهای اقلیمی دما، رطوبت نسبی، سرعت باد و تشعشع خورشیدی جهت پیش‌بینی تبخیر و تعرق یکسال و دو سال آینده استفاده گردید [۳۶]. کسکین و ترزی میزان تبخیر از سطح تشتک را در مقیاس روزانه با استفاده از شبکه‌های عصبی پیش‌بینی کردند [۱۵].

کیسی شبکه‌های عصبی رگرسیونی تعمیم‌یافته را در برآورد تبخیر و تعرق به کار گرفت [۱۸]. زانتی و همکاران و کیم و کیم نیز به کارایی نتایج حاصل از شبکه عصبی مصنوعی در پیش‌بینی و تخمین تبخیر و تعرق اشاره کردند [۱۷، ۴۱]. چاهان و شریاستاوا پیش‌بینی تبخیر و تعرق را با استفاده از سه الگوریتم متفاوت نیوتن، پس انتشار خطا<sup>۴</sup> و لورنبرگ مارکوات<sup>۵</sup> انجام دادند و میزان تبخیر و تعرق ماه آینده را پیش‌بینی کردند. روش محاسبه تبخیر و تعرق در این تحقیق روش پنمن-مانتیس می‌باشد. در نهایت الگوریتم نیوتن در پیش‌بینی تبخیر و تعرق کمترین خطا و بالاترین ضریب کارایی را نشان داد [۸]. نتایج کار تراژور و همکاران نیز در همین راستا گزارش شد [۳۷]. کومار و همکاران نتایج استفاده از تکنیک شبکه عصبی و استنتاج فازی را در پیش‌بینی تبخیر و تعرق با همدیگر مقایسه گردید، نتایج نشان داد هر دو مدل نتایج کارایی مطلوبی داشته‌اند علی‌رغم اینکه کارایی مدل استنتاج فازی به میزان اندکی برتر از شبکه عصبی می‌باشد [۲۱]. کیسی و تمبول پیش‌بینی میزان تبخیر از سطح تشتک را با استفاده از روش فازی ژنتیک<sup>۶</sup> در دو ایستگاه ترکیه انجام دادند و نتایج آن را با روش شبکه عصبی، استنتاج فازی و روش استیفن استیفارد<sup>۷</sup> مقایسه کردند. نتایج برتری روش فازی ژنتیک را در پیش‌بینی‌ها هم در مرحله ارزیابی مدل‌ها و هم در مرحله

<sup>1</sup> Artificial Neural Network

<sup>2</sup> Active Neuro-Fuzzy Inference System

<sup>3</sup> Decision tree

<sup>4</sup> Back propagation

<sup>5</sup> Levenberg-Marquardt

<sup>6</sup> fuzzy genetic

<sup>7</sup> Stephens-Stewart

قابل ملاحظه مدل‌های هوشمند تبخیر تعرق گیاه مرجع می‌باشد [۴۲]. چاری و همکاران، کارایی روش‌های رگرسیونی و شبکه‌های عصبی را با چهار ورودی تبخیر از سطح آزاد آب، دماهای بیشینه و کمینه و رطوبت سطحی در پیش‌بینی تبخیر از سطح ایستابی کم‌عمق برای سه بافت خاک لوم شنی، لومی و لوم رسی و در سه عمق مختلف در پایه‌ی زمانی روزانه مورد بررسی قرار گرفت. نتایج هماهنگی مناسبی را بین مقادیر اندازه‌گیری و پیش‌بینی شده نشان دادند و استفاده صحیح از روش‌های وایازی و شبکه‌های عصبی مصنوعی را در تحلیل سریع تبخیر از خاک مؤثر نشان داد [۷]. نورانی و سیاح‌فرد میزان تبخیر در ایستگاه ارومیه و تبریز را با به کار بردن پارامترهای هواشناسی و استفاده از مدل‌های شبکه‌ی عصبی پرسپترون، المان و شبکه‌های شعاعی انجام دادند و سپس نتایج را با سه مدل فیزیکی تعادل انرژی، ائرودینامیک و مدل پنمن و همچنین مدل رگرسیون چند متغیره مقایسه کردند. سپس آنالیز حساسیت بر روی پارامترهای اقلیمی مؤثر بر تبخیر انجام گرفت [۲۸]. نتایج برتری مدل‌های شبکه‌ی عصبی را نسبت به سایر روش‌ها نشان داد. علاوه بر این عملکرد شبکه‌های پرسپترون نسبت به سایر مدل‌های شبکه‌ی عصبی بهتر می‌باشد و سه پارامتر دما، تشعشع خورشیدی و میزان تبخیر در گام زمانی قبلی، نسبت به سایر پارامترها در پیش‌بینی‌ها مؤثرتر می‌باشد. خوشحال و مکرم، پیش‌بینی تبخیر و تعرق را در منطقه اقلید با استفاده از ۱۰ ترکیب متفاوت از پارامترهای هواشناسی و شبکه‌های پرسپترون انجام دادند. نتایج نشان داد ترکیب پارامترهای دمای بیشینه، دمای کمینه، سرعت باد، رطوبت نسبی و رطوبت مطلق عملکرد بهتری نسبت به سایر ترکیب‌ها در پیش‌بینی تبخیر دارد [۱۶]. پیری و انصاری روش‌های برآورد تبخیر را براساس مدل شبکه‌ی عصبی غیرخطی (NNARX) و تکنیک‌های استنتاج فازی (ANFIS) تطبیقی بررسی کردند. یافته‌ها نشان داد که تکنیک NNARX و ANFIS عملکرد بسیار بهتر از فرمول‌های تجربی دارند. همچنین نویسندگان اذعان نمودند که

تخمین تبخیر نسبت به سه روش دیگر نشان دادند [۱۹]. در ایران نیز تحقیقاتی در راستای برآورد میزان تبخیر صورت گرفته است. شایان نژاد در تحقیقی مقدار تبخیر و تعرق پتانسیل به دست آمده با روش پنمن-مانتیس را با مقدار به دست آمده از روش شبکه‌ی عصبی مصنوعی مقایسه کرد. نتایج نشان داد مقدار خطای به دست آمده از روش پنمن-مانتیس ۱/۲ میلی‌متر در روز و از روش شبکه‌های عصبی مصنوعی ۰/۷ میلی‌متر در روز می‌باشد، بنابراین روش شبکه‌های عصبی مصنوعی در تعیین تبخیر و تعرق پتانسیل بسیار دقیق‌تر از روش پنمن-مانتیس است [۳۱]. مقدم نیا و همکاران روش‌های برآورد تبخیر بر اساس شبکه‌های عصبی مصنوعی (ANN) و تکنیک‌های استنتاج فازی تطبیقی (ANFIS) مورد بررسی قرار دادند. یافته‌ها نشان داد که تکنیک‌های ANN و ANFIS عملکرد بسیار بهتر از فرمول‌های تجربی دارند. از طرفی بین ANN و ANFIS، مدل ANN کمی بهتر است هرچند تفاوت بین دو مدل ناچیز است [۲۳]. صیادی و همکاران عملکرد دو شبکه‌ی MLP و RBF را در برآورد تبخیر و تعرق گیاه مرجع با هم مقایسه کردند. نتایج حاصل نشان داد که شبکه‌های MLP نسبت به RBF در تخمین تبخیر و تعرق گیاه مرجع از دقت نسبتاً بیشتری برخوردارند و تنها مزیت شبکه‌های RBF زمان کمتر مورد نیاز برای آموزش است [۲۸]. زارع و همکاران از شبکه‌ی عصبی مصنوعی و سیستم استنتاج فازی به منظور استفاده از حداقل پارامترهای هواشناسی در پیش‌بینی تبخیر تعرق گیاه مرجع در منطقه همدان استفاده کردند. بدین منظور، با استفاده از آزمون پیرسون، شش پارامتر هواشناسی مورد نیاز در روش پنمن-مانتیس فائو که شامل دمای حداکثر و حداقل، مقادیر رطوبت نسبی حداقل و حداکثر، سرعت باد در ارتفاع دو متری و ساعت آفتابی روزانه می‌باشد، انتخاب گردید. در مجموع بین نتایج شبکه‌ی عصبی مصنوعی و سیستم استنتاج فازی از نظر آماره‌های خطاسنجی اختلاف قابل ملاحظه‌ای مشاهده نشد. همچنین برآوردهای شبکه‌های عصبی با روش استاندارد پنمن-مانتیس نیز نشان دهنده دقت و توانمندی

درخت تصمیم در پیش‌بینی میزان تبخیر ایستگاه سینوپتیک یزد با هم مقایسه گردد و سپس مؤثرترین پارامترها در پیش‌بینی تبخیر شناسایی گردد.

## ۲. روش شناسی

### ۱.۲. داده‌های مورد استفاده

در این تحقیق از آمار ماهانه ایستگاه سینوپتیک یزد در طول سال‌های ۱۹۷۵ لغایت ۲۰۰۴ استفاده شده است. این ایستگاه واقع در شهرستان یزد از سال ۱۳۴۰ شروع به کار کرده و در ارتفاع ۱۲۳۰ متری از سطح دریا و طول جغرافیایی ۱۷° ۵۴ و عرض جغرافیایی ۳۱° ۵۳ قرار دارد. شایان ذکر است به دلیل عدم وجود داده‌های تبخیر در بازه سال‌های ۱۹۷۷ تا ۱۹۸۲ دو بازه زمانی برای متغیرهای هواشناسی در نظر گرفته شده است. جدول ۱ پارامترهای ورودی به مدل‌ها را در این تحقیق نشان می‌دهد. با توجه به جدول ارائه شده ترکیبات متفاوتی از چندین پارامتر ذکر شده ایجاد گردید و تأثیر پیش‌پردازش بر روی این ترکیبات با مدل‌های مختلف مورد بررسی قرار گرفت. در این مرحله با اضافه شدن هر پارامتر، در ترکیبات مختلف، تأثیر اضافه شدن هر پارامتر بر روی خروجی نیز مشخص می‌گردد.

اگرچه روش‌های NNARX و ANFIS قدرتمند هستند، فرایند انتخاب ورودی داده‌ها بسیار پیچیده است [۲۶]. شیری و همکاران با استفاده از مدل جهانی فازی<sup>۱</sup> به پیش‌بینی میزان تبخیر در دو منطقه مرطوب و غیر مرطوب پرداختند. بر این اساس با استفاده از دو ایستگاه در اسپانیا که در دو منطقه مرطوب و غیرمرطوب واقع شده بودند شبکه آموزش داده شد و مورد ارزیابی قرار گرفت سپس با استفاده از مدل توسعه یافته موجود میزان تبخیر با استفاده از پارامترهای هواشناسی در دو منطقه مرطوب و غیرمرطوب ایران تخمین زده شد. نتایج نشان داد استفاده از مدل جهانی فازی به خصوص در مناطق مرطوب نتایج درخشانی را به دنبال دارد [۳۲]. پیری و همکاران دو مدل هیبریدی ANN-COA و ANFIS-COA توسعه دادند و عملکرد آن‌ها با ANN و ANFIS مقایسه می‌شود. نتایج مدل‌ها نشان داد که ترکیب COA با تکنیک‌های ANN و ANFIS دقت مدل‌های ANN و ANFIS را در تمام ایستگاه‌های مورد نظر به میزان قابل توجهی افزایش نمی‌دهد و ترکیب الگوریتم COA با ANN و ANFIS نمی‌تواند یک گزینه مناسب برای برآورد تبخیر روزانه باشد [۲۷]. در تحقیق حاضر سعی شده با به کار بردن برخی از مهم‌ترین پارامترهای هواشناسی مؤثر، به طور هم‌زمان کارایی سه روش داده کاوی شبکه عصبی مصنوعی، استنتاج فازی و

جدول ۱. پارامترهای ورودی مدل‌ها (ایستگاه یزد، سال‌های ۱۹۷۵-۱۹۷۷ و ۱۹۸۲-۲۰۰۴)

علامت اختصاری	پارامتر ورودی	مقادیر کمینه	مقادیر متوسط	مقادیر بیشینه
Tmax	دمای بیشینه (درجه سانتی‌گراد)	۴۲/۶	۲۶/۷	۵/۸
Tmin	دمای کمینه (درجه سانتی‌گراد)	۲۸/۳	۱۲/۲	۳/۴
Tave	دمای متوسط (درجه سانتی‌گراد)	۳۵/۵	۱۹/۴	۶/۱
S	ساعات آفتابی (ساعت)	۱۳۱/۲	۲۷۰/۱	۳۷۶/۸
W <sub>s</sub>	سرعت باد (متر بر ثانیه)	۱/۶	۴/۶	۷/۷
W <sub>a</sub>	جهت باد (درجه)	۲۰	۲۵۰	۳۶۰
R <sub>H</sub>	رطوبت نسبی (درصد)	۸	۲۵/۱	۷۴
E <sub>R</sub>	تبخیر با اعمال ضریب اصلاحی تشتک (میلی‌متر)	۰	۲۵۱	۶۲۴/۷

<sup>۱</sup> Global Neuro Fuzzy

۲۰۰۴)، پیش‌بینی در لایه خروجی (سال ۲۰۰۵) انجام گرفت و با مقادیر واقعی در این سال مقایسه گردید.

## ۲.۲. روش مورد استفاده

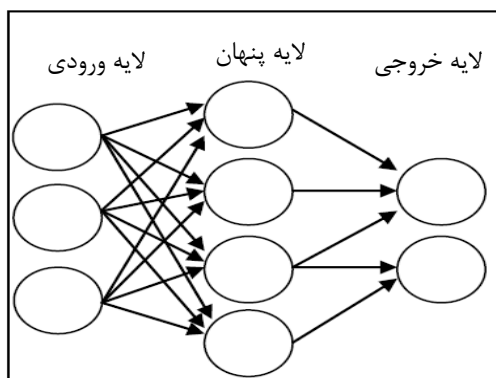
در این تحقیق به منظور پیش‌بینی تبخیر از سه روش مهم داده‌کاوی تحت عنوان شبکه عصبی مصنوعی، شبکه‌های استنتاج فازی و درخت تصمیم رگرسیونی استفاده و شبیه‌سازی‌ها در دو حالت کلی شامل استفاده از داده‌های خام و استفاده از داده‌های نرمال شده (مطابق با رابطه ۱) انجام شد. در این تحقیق سناریو استفاده از داده‌های خام به منظور بررسی تأثیر پیش‌پردازش داده‌ها در مدل‌های داده‌کاوی می‌باشد تا در نهایت مشخص گردد که در پیش‌بینی تبخیر، نرمال‌سازی داده‌ها قبل از ورود به مدل تا چه اندازه در کارایی مدل مؤثر خواهد بود.

$$\frac{X_o - X_{min}}{X_{max} - X_{min}} \quad (1)$$

## ۳.۲. شبکه‌های عصبی پرسپترون چند لایه<sup>۱</sup>

شبکه‌های عصبی پرسپترون انواعی از شبکه‌های پیشخور<sup>۲</sup> هستند که از سه لایه ورودی، لایه پنهان و لایه خروجی تشکیل شده‌اند. شکل ۱ نمونه ساده‌ای از این شبکه‌ها را نشان می‌دهد.

در ابتدا آمار و اطلاعات مربوط به ایستگاه مزبور اخذ و از طریق آزمون توالی، همگنی داده‌ها بررسی و سپس نواقص آماری برطرف گردید. پس از آن، داده‌ها به نسبت ۷۰ به ۳۰ به دو دسته آموزش و آزمون تقسیم شدند و اعتبار سنجی مدل‌ها در مرحله نهایی انجام گرفت. شایان ذکر است نسبت بین داده‌های آموزش و آزمایش می‌تواند متغیر باشد. رایج‌ترین نسبت که به طور معمول در بیشتر پیش‌بینی‌ها صورت می‌گیرد نسبت در نظر گرفته شده در بالا می‌باشد [۱۰، ۲۸، ۴۲]. انتخاب داده‌های آموزش و آزمون به صورت کاملاً تصادفی و توسط نرم‌افزار صورت گرفت. مدل بین‌ورودی‌ها که متغیرهای هواشناسی می‌باشد و خروجی که تبخیر با یک شیفیت ۱۲ ماه بعد می‌باشد آموزش می‌بیند تا رابطه‌ای منطقی را بین ورودی‌ها و خروجی‌ها پیدا کند، سپس در مرحله آزمایش، نتایج حاصل از کارایی مدل‌ها مورد آزمایش قرار گرفته و با یکدیگر مقایسه می‌گردد و مدل برتر شناسایی می‌شود. در نهایت بعد از شناسایی مدل برتر به منظور اعتبار سنجی مدل، داده‌های مربوط به سال ۲۰۰۴ به مدل معرفی و مقادیر تبخیر مربوط به سال ۲۰۰۵ پیش‌بینی گردیدند. تفاوت این مرحله با مرحله قبل در این است که در این مرحله داده‌های سال ۲۰۰۴ و ۲۰۰۵ در مرحله آزمایش و آزمون مورد استفاده قرار نگرفته بودند و بر خلاف مرحله آزمایش و آزمون که در مدل هم ورودی و هم خروجی وارد می‌گردد، در این مرحله تنها با اعمال ورودی‌ها (سال



شکل ۱. نمونه‌ای از شبکه‌های پرسپترون چند لایه استفاده شده در این تحقیق

<sup>۱</sup> Multi layer perceptron

<sup>۲</sup> Feed forward

همگرایی استفاده می‌گردد. برای هر دور اگر کارایی به سمت تابع هدف کاهش پیدا کند، نرخ یادگیری با افزایش فاکتور یادگیری کاهش یافته و چنانچه کارایی افزایش یابد، نرخ یادگیری با کاهش فاکتور یادگیری تعدیل می‌گردد.

## ۲، ۴. شبکه‌های استنتاج فازی

شبکه‌های استنتاج فازی تطبیقی بر اساس الگوریتم فازی سوگنو-تاکاگی<sup>۵</sup> بنا شده است (۱۴). شبکه‌های استنتاج فازی ترکیبی از شبکه‌های عصبی و فازی می‌باشد که در یک چارچوب واحد قرار گرفته است. این شبکه‌ها از پنج لایه تشکیل شده است که شامل لایه‌های ورودی و خروجی می‌باشد (شکل ۲). سیستم فازی نوع سوگنو با شبکه‌های عصبی تطبیقی ترکیب شده است که در این مطالعه برای پیش‌بینی تبخیر استفاده گردیده است و آموزش‌ها بر اساس الگوریتم یادگیری پس انتشار خطا طرح ریزی شده است [۳۵]. در شکل ۲، X و Y مقادیر ورودی هستند که به نرون‌های لایه ورودی (لایه اول) وارد می‌شود، سپس این مقادیر وارد توابع عضویت زنگوله‌ای<sup>۶</sup> شکل می‌شود و سپس مقادیر خروجی حاصل می‌گردد. شکل ۲ نمونه‌ای از شبکه‌های استنتاج فازی و لایه‌های آن را نشان می‌دهد.

شبکه‌های پرسپترون با الگوریتم‌های متفاوتی آموزش می‌یابند. در این مطالعه الگوریتم یادگیری لورنبرگ مارکوات استفاده گردیده است که بر اساس قانون دلتا<sup>۱</sup> پایه‌گذاری شده است. در این قانون مقدار تصحیح اعمال شده به وزن‌های ورودی یک نرون در هر تکرار، تابعی از ورودی خالص نرون، میزان خطای شبکه و نرخ یادگیری شبکه می‌باشد.

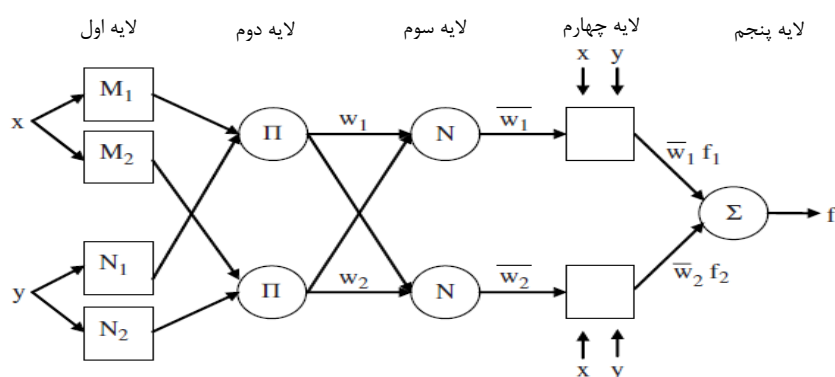
چنانچه  $x_i = (i=1,2,\dots,n)$  به عنوان ورودی و  $w_i (i=1,2,\dots,n)$  وزن هر ورودی باشد آنگاه ورودی هر شبکه برای هر گره طبق معادله ۲ بیان می‌گردد.

$$net = \sum_{i=1}^n x_i w_i \quad (2)$$

که ورودی هر گره هنگامی که از تابع انتقال<sup>۲</sup> عبور می‌کند آنگاه خروجی هر گره به صورت  $y = f(x)$  محاسبه می‌گردد. شایان ذکر است در این تحقیق از تابع انتقال و غیرخطی سیگموئید<sup>۳</sup> همانند زیر استفاده می‌شود

$$y = f(net) = \frac{1}{1 + e^{-net}} \quad (3)$$

در اغلب پیش‌بینی‌هایی که توسط شبکه عصبی صورت می‌گیرد نرخ یادگیری تطبیقی<sup>۴</sup> برای افزایش سرعت



شکل ۲. ساختار ساده‌ای از شبکه‌های استنتاج فازی مورد استفاده در این تحقیق

<sup>1</sup> Delta

<sup>2</sup> Translate Function

<sup>3</sup> Sigmuid

<sup>4</sup> Adaptive learning rate

<sup>5</sup> Tagaki-Sugeno-Kang(TSK)

<sup>6</sup> Bell

## ۵.۲. درخت تصمیم رگرسیونی

مدل نیمه هوشمند درخت تصمیم رگرسیونی یکی از روش های داده کاوی است. این مدل می تواند همه یا تعدادی از متغیرهای مستقل در اختیار را برای پیش بینی مورد استفاده قرار دهد. معیار مورد استفاده در درخت تصمیم رگرسیونی، انحراف حداقل مربعات<sup>۲</sup> نام دارد (۵). این معیار به صورت زیر تعریف می شود:

$$SS(t) = \sum_{i=1}^{N_t} (y_i(t) - \bar{y}(t))^2 \quad (۸)$$

$N_t$ : تعداد رکوردها (داده ها) در گره برگ  $t$ .

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

$\bar{Y}(t)$ : میانگین مقادیر متغیر هدف برای همه گره ها.

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

$$Q(s,t) = SS(t) - SS(t_R) - SS(t_L) \quad (۹)$$

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

## ۶.۲. معیارهای ارزیابی کارایی مدلها

به منظور ارزیابی کارایی مدلها در پیش بینی مقدار تبخیر از آماره ضریب همبستگی<sup>۳</sup> استفاده شد. همچنین جهت برآورد خطا از دو آماره ریشه میانگین مربعات خطا<sup>۴</sup>، میانگین مطلق خطا<sup>۵</sup> و شاخص ناش-ساتکلیف<sup>۶</sup> استفاده شد (ضریب ME بین منفی بی نهایت تا یک تغییر می کند که این مقدار هر چه به عدد یک نزدیک تر باشد برآوردها دقیق تر هست). معادلات آماره های فوق در معادلات ۱۰ تا ۱۳ آورده شده است.

$W$  خروجی یک نرون هست و  $\mu$  تابع عضویت می باشد.  $iM$  و  $iN$  به ترتیب مجموعه های فازی نرون های  $x$  و  $y$  هستند. تابع عضویت مجموعه های فازی<sup>۱</sup> معمولاً توسط توابع زنگی شکل تعمیم یافته معادل رابطه ۴ تعریف می شود.

$$\mu_{M_i}(x) = \frac{1}{1 + \left| \frac{x - c_i}{a_i} \right|^{2N_i}} \quad (۴)$$

پارامترهای  $a_i$ ،  $b_i$  و  $c_i$  پارامترهای توابع عضویت این مجموعه هستند. این پارامترها به نام پارامترهای شرط نیز نامیده می شوند. سپس در لایه دوم تابع عضویت  $x$  در تابع عضویت  $y$  ضرب مطابق با معادله ۵ محاسبه می شود

$$w_i = \mu_{M_i}(x) \mu_{N_i}(y), (i = 1, 2) \quad (۵)$$

در لایه سوم نسبت خروجی هر نرون به خروجی کلیه نرون ها طبق معادله ۶ محاسبه می گردد

$$\bar{w}_i = \frac{w_i}{\sum w_i}, (i = 1, 2) \quad (۶)$$

نرون ها در لایه چهارم از نرخ حاصل از لایه سوم به عنوان وزن استفاده می کنند و در پارامتر خروجی ضرب می شود. در این لایه قانون اگر-بنابراین حاکم است. به عنوان مثال اگر  $x$  معادل  $iM$  باشد و  $y$  برابر  $iN$  باشد پس اگر  $f_i = (p_i x + q_i y + r_i)$  بنابراین

$$\bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (۷)$$

در رابطه بالا  $r_i$ ،  $p_i$  و  $q_i$  پارامترهای تطبیقی لایه اند و به نام پارامترهای نتیجه خوانده می شوند. در نهایت خروجی شبکه  $f$  به وسیله نرون های پنجمین لایه حاصل می گردد که به عنوان مجموع کل سیگنال خروجی محسوب می گردد که طبق معادله ۶ محاسبه می گردد. بدین صورت می توان یک سیستم فازی که قابلیت یادگیری دارد را پیاده سازی کرد (۳).

<sup>1</sup> Fuzzy set

<sup>2</sup> Least Square Deviation

<sup>3</sup> correlation coefficient(r)

<sup>4</sup> Root Mean Squar Error(RMSE)

<sup>5</sup> Mean Absolute Error(MAR)

<sup>6</sup> Nash & Sutcliffe

مرحله آزمایش و در حالت استفاده از داده های خام در جدول ۲ ارائه شده است. هریک از پارامترهای ترکیب ورودی در جدول ۱ مشخص گردیده است.

جدول ۲ نشان می دهد برترین مدل شبکه عصبی مصنوعی در حالت استفاده از داده های خام، ترکیب "میانگین دما، متوسط بیشینه دما، تبخیر، میانگین رطوبت نسبی" می باشد. این مدل می تواند با دقت مناسبی به پیش بینی میزان تبخیر سال بعد پردازد. نتایج حاصل از اجرای مدل شبکه عصبی مصنوعی در حالت نرمال در جدول ۳ ارائه شده است.

نتایج نشان می دهد شبکه عصبی مصنوعی در حالت استفاده از داده های نرمال با استفاده از ترکیب ردیف دوم "میانگین دما، متوسط کمینه دما، متوسط بیشینه دما، تبخیر، میانگین رطوبت نسبی، سرعت باد، ساعات آفتابی" به بهترین نتایج دست یافته است. این مدل نیز می تواند با دقت مناسبی به پیش بینی میزان تبخیر سال بعد پردازد (جدول ۳).

در مجموع مدل شبکه عصبی مصنوعی در حالت استفاده از داده های نرمال میزان خطای کمتری داشته است و نتایج به میزان ۱/۸ درصد بهبود یافته است.

$$r = \frac{\sum_i (O_i - \bar{O}_i)(P_i - \bar{P}_i)}{N} \quad (10)$$

$$= \frac{\sum_i (P_i - \bar{P}_i)^2}{\sqrt{\sum_i (P_i - \bar{P}_i)^2} \sqrt{\sum_i (O_i - \bar{O}_i)^2}}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2 \quad (11)$$

$$RMSE = \sqrt{\left(\frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2\right)} \quad (12)$$

$$ME = 1 - \left[ \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O}_i)^2} \right] \quad (13)$$

که در آن  $O_i$  داده های مشاهداتی  $P_i$  داده های شبیه سازی شده و  $n$  تعداد داده ها است.

### ۳. نتایج

#### ۱،۳. نتایج حاصل از اجرای مدل شبکه عصبی

##### مصنوعی

نتایج حاصل از اجرای مدل شبکه عصبی مصنوعی در

جدول ۲. نتایج حاصل از ارزیابی شبکه های عصبی مصنوعی در حالت استفاده از داده های اصلی

مرحله	ترکیب ورودی	تابع میانی	تابع خروجی	R Evaluation	RMSE	MAE	ME
آموزش	$T_{ave} - T_{max} - E_R$	Sigmoid	Sigmoid	۰/۹۶	۵/۰۷	۳۵/۳۵	-۰/۴۱
آزمایش				۰/۹۵	۵/۸	۳۷/۳۴	-۰/۲۴
آموزش	$T_{ave} - T_{max} - E_R - R_H$	Sigmoid	Sigmoid	۰/۹۶	۵/۵	۳۶/۴۲	-۰/۳۵
آزمایش				۰/۹۶	۵/۴	۳۶/۷۲	-۰/۰۵
آموزش	$T_{ave} - T_{max} - E_R - R_H - W_s$	Sigmoid	Sigmoid	۰/۹۶	۶/۶۷	۴۲/۴۳	-۰/۱۷
آزمایش				۰/۹۵	۶/۶۲	۴۱/۵۹	-۰/۰۶
آموزش	$T_{ave} - T_{max} - E_R - R_H - W_s - S$	Sigmoid	Sigmoid	۰/۹۷	۴/۴۶	۴۲/۷۸	-۰/۴۶
آزمایش				۰/۹۵	۵/۷۸	۴۰/۰۷	-۰/۴۰
آموزش	$T_{ave} - T_{min} - T_{max} - E_R - R_H - W_s - S$	Sigmoid	Sigmoid	۰/۹۷	۵/۶۰	۳۸/۳	-۰/۸۹
آزمایش				۰/۹۶	۶/۲۳	۴۱/۴۳	-۰/۵۵



جدول ۳. نتایج حاصل از ارزیابی شبکه‌های عصبی مصنوعی در حالت استفاده از داده‌های نرمال شده

مرحله	ترکیب ورودی	تابع میانی	تابع خروجی	R evaluation	RMSE	MAE	ME
آموزش	$T_{ave} - T_{max} - E_R$	Sigmoid	Sigmoid	۰/۹۴	۶/۶	۴۲/۴	۰/۲۵
آزمایش				۰/۹۰	۸/۶۱	۴۸/۰۱	-۰/۱۸
آموزش	$T_{ave} - T_{max} - E_R - R_H$	Sigmoid	Sigmoid	۰/۹۶	۵/۰۱	۳۵/۶۷	۰/۱۱
آزمایش				۰/۹۵	۶/۱	۳۹/۰۰	۰/۳
آموزش	$T_{ave} - T_{max} - E_R - R_H - W_s$	Sigmoid	Sigmoid	۰/۹۶	۵/۱	۳۲/۱	۰/۳۵
آزمایش				۰/۹۶	۵/۳	۳۶/۳	۰/۴۸
آموزش	$T_{ave} - T_{max} - E_R - R_H - W_s - S$	Sigmoid	Sigmoid	۰/۹۵	۶/۱۱	۴۰/۰۱	-۰/۱۱
آزمایش				۰/۹۶	۶/۲۷	۴۲/۰۳	-۰/۰۴
آموزش	$T_{ave} - T_{min} - T_{max} - E_R - R_H - W_s - S$	Sigmoid	Sigmoid	۰/۹۶	۴/۴۲	۳۰/۴	-۰/۲۱
آزمایش				۰/۹۷	۵/۱	۳۶/۳	-۰/۴۸

در حالت استفاده از داده‌های خام در جدول ۴ ارائه شده است.

### ۲،۳. نتایج حاصل از اجرای مدل شبکه‌های

### استنتاج فازی

نتایج حاصل از اجرای مدل شبکه‌های استنتاج فازی

جدول ۴. نتایج حاصل از ارزیابی شبکه‌های استنتاج فازی در حالت استفاده از داده‌های اصلی

مرحله	ترکیب ورودی	تابع عضویت	مدل فازی	R Evaluation	RMSE	MAE	ME
آموزش	$T_{ave} - T_{max} - E_R$	Bell	TSK	۰/۹۷	۳/۲	۳۷/۸	-۰/۱۴
آزمایش				۰/۹۶	۵/۴	۳۹	-۰/۲۴
آموزش	$T_{ave} - T_{max} - E_R - R_H$	Bell	TSK	۰/۹۶	۴/۸	۴۰/۰۳	-۰/۷۱
آزمایش				۰/۹۶	۵/۹	۴۲/۲۹	-۰/۷۸
آموزش	$T_{ave} - T_{max} - E_R - R_H - W_s$	Bell	TSK	۰/۹۵	۳/۴۱	۳۷/۰۱	-۰/۲۱
آزمایش				۰/۹۶	۵/۹۹	۳۸/۰۹	-۰/۰۲
آموزش	$T_{ave} - T_{max} - E_R - R_H - W_s - S$	Bell	TSK	۰/۹۵	۴/۴۵	۳۸/۱۱	-۰/۱۵
آزمایش				۰/۹۶	۵/۸	۴۰/۲۶	-۰/۰۹
آموزش	$T_{ave} - T_{min} - T_{max} - E_R - R_H - W_s - S$	Bell	TSK	۰/۹۶	۵/۶	۴۰/۳۵	-۰/۵۶
آزمایش				۰/۹۶	۵/۵	۴۰/۱۳	-۰/۳۶

نتایج حاصل از اجرای مدل شبکه‌های استنتاج فازی در حالت استفاده از داده‌های نرمال در جدول ۵ ارائه شده است. مدل شبکه‌های استنتاج فازی نیز با استفاده از داده‌های نرمال به نتایج بهتری در پیش‌بینی مقادیر تبخیر سال بعد دست یافته است. بهبود نتایج در این حالت معادل ۲/۵ درصد می‌باشد.

نتایج نشان می‌دهد کلیه مدل‌سازی‌های صورت گرفته با شبکه‌های استنتاج فازی جهت پیش‌بینی تبخیر با استفاده از داده‌های خام موفق بوده است، اما ترکیب "میانگین دما، متوسط بیشینه دما، تبخیر، میانگین رطوبت نسبی، سرعت باد" به عنوان ترکیب برتر شناخته شد (جدول ۴).

جدول ۵. نتایج حاصل از ارزیابی شبکه‌های استنتاج فازی در حالت استفاده از داده‌های نرمال شده

مرحله	ترکیب ورودی	تابع عضویت	مدل فازی	R evaluation	RMSE	MAE	ME
آموزش	$T_{ave} - T_{max} - E_R$	Bell	TSK	۰/۹۵	۶/۴	۴۵/۲	-۰/۱۹
آزمایش				۰/۹۱	۸/۵	۴۸/۷	-۰/۲۱
آموزش	$T_{ave} - T_{max} - E_R - R_H$	Bell	TSK	۰/۹۵	۳/۲	۳۱/۷	۰/۴۵
آزمایش				۰/۹۵	۵/۹	۳۸/۸	۰/۱۴
آموزش	$T_{ave} - T_{max} - E_R - R_H - W_s$	Bell	TSK	۰/۹۵	۴/۴	۳۴/۳	۰/۰۹
آزمایش				۰/۹۶	۵/۳	۳۸/۱	۰/۲۴
آموزش	$T_{ave} - T_{max} - E_R - R_H - W_s - S$	Bell	TSK	۰/۹۶	۴/۹	۳۷/۸	۰/۰۷
آزمایش				۰/۹۶	۵/۴	۳۷/۱	۰/۱۱
آموزش	$T_{ave} - T_{min} - T_{max} - E_R - R_H - W_s - S$	Bell	TSK	۰/۹۶	۳/۳	۳۶/۷	-۰/۰۱
آزمایش				۰/۹۶	۵/۳	۳۸/۲	۰/۰۷

ترکیب "میانگین دما، متوسط کمینه دما، متوسط بیشینه دما، تبخیر، میانگین رطوبت نسبی، سرعت باد، ساعات آفتابی" به عنوان ترکیب برتر شناخته شد (جدول ۶). همچنین نتایج حاصل از اجرای مدل درخت تصمیم رگرسیونی در حالت استفاده از داده‌های نرمال در جدول ۷ ارائه شده است.

### ۳.۳. نتایج حاصل از اجرای مدل درخت تصمیم

#### رگرسیونی

نتایج حاصل از اجرای مدل درخت تصمیم رگرسیونی در حالت استفاده از داده‌های خام در جدول ۶ ارائه شده است. در شبیه‌سازی‌های انجام شده با مدل درخت تصمیم رگرسیونی در حالت استفاده از داده‌های خام

جدول ۶. نتایج حاصل از ارزیابی مدل درخت تصمیم رگرسیونی در حالت استفاده از داده‌های اصلی

ترکیب ورودی	تعداد سطح	تعداد قانون	R evaluation	RMSE	MAE	ME
$T_{ave} - T_{max} - E_R$	۳۲	۴۵	۰/۹۴	۶۰/۷۳	۴۸/۴۴	-۳/۲۷
$T_{ave} - T_{max} - E_R - R_H$	۳۵	۴۳	۰/۹۳	۶۵/۵۵	۵۱/۲۳	-۳/۷۶
$T_{ave} - T_{max} - E_R - R_H - W_s$	۳۴	۴۳	۰/۹۲	۶۷/۰۶	۵۴/۴۲	-۴/۴
$T_{ave} - T_{max} - E_R - R_H - W_s - S$	۴۲	۴۸	۰/۹۶	۵۳/۱۱	۴۲/۲۷	-۳/۳
$T_{ave} - T_{min} - T_{max} - E_R - R_H - W_s - S$	۳۱	۴۰	۰/۹۶	۴۷/۴۷	۳۸/۸۳	-۲/۹۱

جدول ۷. نتایج حاصل از ارزیابی مدل درخت تصمیم رگرسیونی در حالت استفاده از داده‌های نرمال

ترکیب ورودی	تعداد سطح	تعداد قانون	R evaluation	RMSE	MAE	ME
$T_{ave} - T_{max} - E_R$	۳۲	۴۵	۰/۹۴	۶۰/۷۳	۴۸/۴۴	-۳/۲۷
$T_{ave} - T_{max} - E_R - R_H$	۳۵	۴۳	۰/۹۳	۶۵/۵۵	۵۱/۲۳	-۳/۷۶
$T_{ave} - T_{max} - E_R - R_H - W_s$	۳۴	۴۳	۰/۹۲	۶۷/۰۶	۵۴/۴۲	-۴/۴
$T_{ave} - T_{max} - E_R - R_H - W_s - S$	۴۲	۴۸	۰/۹۵	۵۳/۱۱	۴۲/۲۷	-۳/۳
$T_{ave} - T_{min} - T_{max} - E_R - R_H - W_s - S$	۳۱	۴۰	۰/۹۶	۴۷/۴۷	۳۸/۸۳	-۲/۹۱

بدون دخالت کاربر، ورودی‌های مهم‌تر را برای ایجاد قوانین پیش‌بینی استفاده و ورودی‌های ضعیف‌تر را حذف نماید. در تمامی مدل‌سازی صورت گرفته در این تحقیق با درخت تصمیم، مدل تمامی پارامترهای مستقل معرفی شده را به کار برد. این نشان‌دهنده تأثیر تمامی متغیرهای مستقل "متوسط کمینه دما، متوسط بیشینه دما، میانگین دما، میانگین رطوبت نسبی، بارش، سرعت باد و تبخیر" در پیش‌بینی میزان تبخیر سال بعد می‌باشد (جدول‌های ۶ و ۷).

جدول ۷ نشان می‌دهد که نرمال کردن داده‌ها هیچ تأثیری در بهبود نتایج مدل درخت تصمیم رگرسیونی نداشته است. به نظر می‌رسد این موضوع ناشی از نحوه کارکرد مدل بوده و عملیات نرمال‌سازی که داده‌ها را در دامنه عددی ۱-۰ محدود می‌سازد، بر کارایی درخت تأثیری ندارد. این موضوع با یافته‌های دستورانی و همکاران مبنی بر عدم تأثیر نرمال‌سازی داده‌ها بر کارایی مدل درخت تصمیم رگرسیونی در پیش‌بینی مقادیر بارش مطابقت دارد [۱۰]. همچنین مدل درخت تصمیم قادر است

جدول ۸. کیفیت نتایج ترکیبات برتر در این تحقیق با سه مدل شبکه‌ی عصبی مصنوعی، شبکه‌های استنتاج فازی و درخت تصمیم رگرسیونی

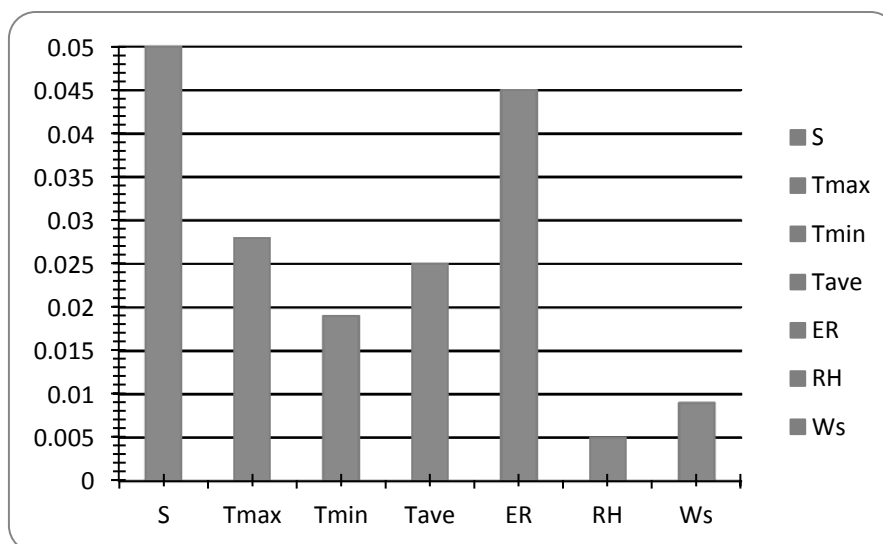
ME	MAE	RMSE	R evaluation	نوع داده	مدل	ترکیب ورودی
-۰/۴۸	۳۶/۳	۵/۱	۰/۹۷	نرمال	شبکه‌ی عصبی مصنوعی	Tave - Tmin -Tmax - ER - RH - Ws -S
۰/۱۱	۳۷/۱	۵/۴	۰/۹۶	نرمال	شبکه‌های استنتاج فازی	Tave -Tmax - ER - RH - Ws -S
-۲/۹۱	۳۸/۸	۴۷/۵	۰/۹۶	خام یا نرمال	مدل درخت تصمیم رگرسیونی	Tave - Tmin -Tmax - ER - RH - Ws -S

[۱۱،۳۴]. نتایج تحقیق نورانی و سیاح فرد نیز سه پارامتر "دما، تشعشع خورشیدی و میزان تبخیر" در گام زمانی قبلی را نسبت به سایر پارامترها در پیش‌بینی پدیده تبخیر مؤثرتر دانست [۳۳]. سین پارامترهای "دما، رطوبت، فشار بایومتریک و سرعت باد" را عامل‌های کنترل‌کننده و حائز اهمیت در زمینه تبخیر می‌داند [۲۵]. گوپتا، نیز به اهمیت تشعشعات خورشیدی، رطوبت نسبی، سرعت باد و دمای سطح آب و اتمسفر به عنوان کلیدی‌ترین فاکتورها در بحث تبخیر اشاره کرده است [۱۲]. بدیهی است اولویت بندی و شناخت فاکتورهای مؤثر در مدل‌سازی یک پدیده به شرایط منطقه مورد مطالعه بستگی دارد و غالبیت تأثیر یک یا چند پارامتر در منطقه‌ای به منطقه دیگر متفاوت خواهد بود. این موضوع که مبنای تحلیل‌های منطقه‌ای در بررسی پدیده‌های هیدرولوژیکی است، می‌تواند به شناسایی مناطق همگن در خصوص پدیده تبخیر کمک نموده و در مناطق فاقد آمار تبخیر مورد استفاده قرار گیرد.

کیفیت نتایج ترکیبات برتر شناخته‌شده در این تحقیق نشان می‌دهد هر سه مدل شبکه‌ی عصبی مصنوعی، شبکه‌های استنتاج فازی و مدل درخت تصمیم قادرند با استفاده از برخی پارامترهای اقلیمی به پیش‌بینی مقدار تبخیر ماهانه ۱۲ ماه قبل از وقوع پردازند ولی بر اساس نتایج این تحقیق، از نظر دقت مدل شبکه‌ی عصبی مصنوعی در جایگاه نخست، مدل شبکه‌های استنتاج فازی در جایگاه دوم و مدل درخت تصمیم رگرسیونی در جایگاه سوم قرار گرفت. به عبارتی مدل شبکه‌ی عصبی مصنوعی با استفاده از داده‌های نرمال شده مربوط به پارامترهای "میانگین دما، متوسط کمینه دما، متوسط بیشینه دما، تبخیر، میانگین رطوبت نسبی، سرعت باد، ساعات آفتابی" می‌تواند مقدار تبخیر ماهانه ۱۲ ماه قبل از وقوع را پیش‌بینی نماید (جدول ۸). این در حالی است که در تحقیق انجام شده توسط سادهر و همکاران و ابراهیمیان و همکاران، پارامتر "دما" به عنوان مهم‌ترین پارامتر برای مدل‌سازی و برآورد تبخیر معرفی گردید

ذرات از سطح آب و فرایند تبخیر تأثیر بیشتری دارد [۹]. این موضوع با نتایج به دست آمده از آنالیز حساسیت پارامترهای مؤثر در این تحقیق مبنی بر تأثیر به سزای پارامتر ساعات آفتابی در برآورد تبخیر مطابقت دارد.

در این ارتباط باید گفت نتایج تحقیقات انجام شده توسط سرانو و همکاران، علیزاده و خلیلی، بیات و میرلطیفی نشان داد که ساعات آفتابی بیشترین تأثیر را روی تابش خورشیدی دارد [۲۰۴، ۲۹]. همچنین میزان تابش خورشیدی در بالا بردن میزان مکانیسم جابه‌جایی



شکل ۳. نتایج حاصل از تحلیل حساسیت پارامترهای مؤثر در پیش‌بینی تبخیر

درصد می‌باشد. بدیهی است پارامترهای اقلیمی با ضریب تغییرات کم‌تر دارای رژیم منظم‌تری بوده و مدل‌سازی آن‌ها با خطای کمتری همراه است [۲۲].

در خاتمه به‌منظور اطمینان از کارایی مدل برتر در پیش‌بینی‌ها اقدام به اعتبار‌سنجی مدل گردید. همان‌گونه که قبلاً نیز ذکر گردید در این مرحله برخلاف مرحله آموزش و آزمایش، داده‌های خروجی به مدل وارد نشد و تنها با اعمال ورودی‌ها (سال ۲۰۰۴)، خروجی (تبخیر سال ۲۰۰۵) پیش‌بینی گردید و با مقادیر واقعی آن مقایسه گردید. شایان ذکر است داده‌های سال ۲۰۰۴ و ۲۰۰۵ در هیچ‌کدام از مراحل آموزش و آزمایش استفاده نگردیده است. نتایج مذکور در جدول ۹ ارائه شده است.

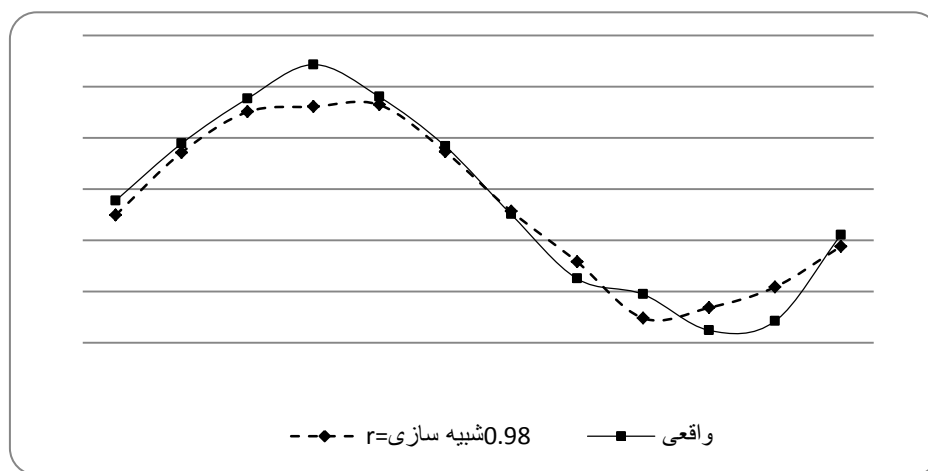
در مجموع ملاحظه می‌شود در شبیه‌سازی‌های صورت گرفته در این تحقیق، تفاوت قابل‌ملاحظه‌ای در زمان استفاده از داده‌های خام و داده‌های نرمال شده وجود ندارد و پردازش داده‌ها تأثیر چندانی در بهبود نتایج مدل نداشته است، این در حالی است که حبیبی پور و همکاران و افخمی و همکاران در تحقیق خود اثر پردازش داده‌های ورودی را در مدل‌سازی بارش ایستگاه یزد مثبت گزارش کرده‌اند [۱۰۱۳]. به نظر می‌رسد تأثیر نداشتن عملیات پردازشی در بهبود کارایی سه مدل داده‌کاوی مورد استفاده در این تحقیق به دلیل ماهیت پارامتر تبخیر و ضریب تغییرات کوچک‌تر این پارامتر در مقایسه با بارش است. به طوری که ضریب تغییرات تبخیر در ایستگاه یزد ۶۷/۴ درصد و ضریب تغییرات بارش ۲۰۳/۶

جدول ۹. نتایج حاصل از پیش‌بینی داده‌های تبخیر با استفاده از مدل برتر (ANN در حالت نرمال)، سال ۲۰۰۵

MAE	RMSE	R evaluation	تابع خروجی	تابع میانی	ترکیب ورودی
۳۳/۳	۱۱/۵۱	۰/۹۸	Sigmoid	Sigmoid	Tave - Tmin - Tmax - E <sub>R</sub> - R <sub>H</sub> - W <sub>s</sub> - S

(ماهه ای سرد) دارای خطای بیشتری است و در بین این دو مقدار (بهار و پاییز) میزان خطا کاهش محسوسی داشته است.

هرچند آماره‌های خطا مؤید کارایی بالای مدل برتر است، ولی شکل ۴ نشان می‌دهد مدل مزبور در پیش‌بینی مقادیر حداکثر تبخیر (ماهه ای گرم) و مقادیر حداقل



شکل ۴. نمودار مقایسه داده‌های پیش‌بینی شده در مقابل مقادیر واقعی با استفاده از مدل برتر (ANN در حالت نرمال)، سال ۲۰۰۵

سه مدل رایج داده‌کاوی (شبکه عصبی مصنوعی، شبکه‌های استنتاج فازی و درخت تصمیم رگرسیون) برای پردازش، شبیه‌سازی، تجزیه و تحلیل و کشف روابط حاکم بر پدیده تبخیر استفاده شد. علیرغم اینکه نوسانات شدید پارامترهای اقلیمی در مناطق خشک، توانایی مدل‌های داده‌کاوی را در یادگیری روابط فی‌مابین این پارامترها تحت‌الشعاع قرار می‌دهد، ولی به دلیل ضریب تغییرات کوچک پارامتر تبخیر در مقایسه با پارامترهایی نظیر بارش؛ مدل‌سازی پدیده تبخیر با موفقیت همراه بود و مدل شبکه عصبی مصنوعی در حالت استفاده از داده‌های نرمال شده، به دلیل خطای کمتر به عنوان مدل برتر معرفی گردید. این مدل قادر است با استفاده از ۷

#### ۴. بحث و نتیجه‌گیری

تبخیر از ویژگی‌های کلیدی اقلیمی در هر محیط محسوب می‌شود. آگاهی از مقدار تبخیر و مدل‌سازی آن به عنوان یکی از متغیرهای مهم هیدرولوژیکی در تحقیقات کشاورزی و حفاظت آب و خاک اهمیت زیادی دارد. این در حالی است که در دهه‌های اخیر تکنیک‌های هوش مصنوعی در تخمین و پیش‌بینی پدیده‌های غیرخطی توانایی و انعطاف‌پذیری بالایی از خود نشان داده‌اند. از طرفی پارامترهای هواشناسی از پایگاه داده مناسبی در ارگان‌های زیربسط برخوردار می‌باشند. از این رو می‌توان از تکنیک‌های داده‌کاوی برای حل مسئله و پیش‌بینی مقادیر تبخیر کمک گرفت. در این تحقیق از

مقایسه، برتری شاخص های خشک سالی نظیر RDI که به پارامتر تبخیر هم توجه دارند نسبت به شاخص های خشک سالی یک پارامتر نظیر SPI که صرفاً پارامتر بارش را در نظر می گیرند در تحقیقات متعددی تأیید شده است (۲۴، ۳۰، ۳۸، ۳۹). علاوه بر آن با شناخت فاکتورهای مؤثر در مدل سازی تبخیر می توان از طریق تحلیل منطقه ای و شناسایی مناطق همگن به پیش بینی تبخیر در مناطق فاقد آمار تبخیر پرداخت.

پارامتر "میانگین دما، متوسط کمینه دما، متوسط بیشینه دما، تبخیر، میانگین رطوبت نسبی، سرعت باد، ساعات آفتابی" به برآورد مقادیر تبخیر ماهانه ۱۲ ماه قبل از وقوع بپردازد. بنابراین با استفاده از روش های داده کاوی می توان به مدل سازی و برآورد پارامتر تبخیر پرداخت. این موضوع می تواند برای پیش بینی وضعیت خشک سالی ۱۲ ماه قبل از وقوع نیز مؤثر واقع شود. چرا که مطالعات و تحقیقات جدید مؤید نقش توأم دو پارامتر بارش و تبخیر در بروز پدیده خشک سالی می باشد. به طوری که در مقام

## References

- [1] Afkhami, H., Dastorani, M. T., Malekinezhad, H. and Mobin, M. H. (2010). Effect of climatic elements in Increasing the accuracy using artificial neural network on forecasting of drought in Yazd. *Journal of Science and Technology of Agriculture and Natural Resources, Water and Soil Science*, 51(14),157-170.
- [2] Alizade, A. and Khalili, N. (2009). Estimation of angstrom coefficient and developing a regression equation for solar radiation estimation (case study: Mashhad). *Journal of Water and Soil*. 23(1), 229-238.
- [3] Azamathulla, H. Md., Chang, C. K., Ghani, A. Ab., Ariffin, J., Azazi Zakaria, N. and Abu Hasan, Z. (2009). An ANFIS-based approach for predicting the bed load for moderately sized rivers. *Journal of Hydro-environment Research*, 3(1),35-44.
- [4] Bayat, K. and Mirlatifi, M. (2009). Estimation of daily solar radiation using regression models and artificial neural network. *Journal of Agricultural Sciences and Natural Resources*.16(3), 270-279.
- [5] Breiman, L., Friedman, J., Olshen, R. and Stone, C. (1984). *Classification and Regression Trees*. Chapman & Hall/CRC Press, Boca Raton, FL.
- [6] Bruton, J. M., McClendon, R. W. and Hoogenboom, G. (2000). Estimating daily panEvaporation with artificial neural network. *Trans, ASAE*, 43(2), 492-4962.
- [7] Chari, M. M., Afrasiab, P., Piri, J. and Delbari, M. (2011). Predicting evaporation from a shallow water table using artificial neural network and simulations of Vayazy. *Journal of Water Engineering*, 8(4),11-20.
- [8] Chauhan, S. and Shrivastava, R. K. (2009). Reference evapotranspiration forecasting using different artificial neural networks algorithms. *Journal of Civil Engineering*, 36 (9),1491-1505.
- [9] Chow, V., Maidment, D. and Mays, L. (1988). *Applied Hydrology*, N.Y.:McGraw-Hill Pub, New York.
- [10] Dastorani, M. T., Habibipur, A., Ekhtesasi, M. R., Talebi, A. and Mahjobi, J. (2012). Evaluation of Performance of the decision tree model on prediction of rainfall. *Journal of Iran-Water Resources Research*,3(8),14.
- [11] Ebrahimian, H., Liaghat, A. and Bazrafshan, M. (2011). Estimation of some climatic factors Using transfer functions. *Journal of Iran-Watershed Management Science and Engineering*, 14(5),77-80.
- [12] Gupta. B. (1992). *Engineering Hydrology*, Jain, India: N.C.
- [13] Habibipur, A., Dastorani, M. T., Ekhtesasi, M. R. and Afkhami, H. (2011). Evaluation of the Effects of Data range Modification on Efficiency of Regression Decision Tree and Artificial Neural Networks for Drought Prediction. *Journal of Watershed Management Research*, 2(3), 63-79.

- [14] Jang, J. S. R., Sun, C. T. and Mizutani, E. (1997). *Neuro-Fuzzy and Soft Computing, A Computational Approach to Learning and Machine Intelligence*, Prentice Hall, NJ, USA ISBN. 0-13-261066-3.
- [15] Keskin, M. E. and Terzi, O. (2006). Artificial neural networks models of daily pan evaporation. *Journal of Hydrological Engineering*, 11(1), 65–70.
- [16] khoshhal, J. and mokarram, M. (2012). Model for Prediction of Evapotranspiration Using MLP Neural Network. *Journal of Enviromental Sciences*, Volume 3( 3),1000-1009.
- [17] Kim,S. and Kim, H.S. (2008). Neural networks and genetic algorithm approach for nonlinear evaporation and evapotranspiration modeling. *Journal of Hydrology*, 351(3-4), 299–317.
- [18] Kisi, O. (2006). Generalized regression neural networks for evapotranspiration modeling. *Journal of Hydrological Sciences*, 51(6), 1092–1105.
- [19] Kisi, O. and Tombul, M.(2013). Modeling monthly pan evaporations using fuzzy genetic approach. *Journal of Hydrology*, 477(16), 203–212.
- [20] Kumar, M., Raghuwanshi, N., Singh, R., Wallender, W. and Pruitt, W. (2002). Estimating Evapotranspiration using Artificial Neural Network. *Journal of Irrigation and Drainage Engineering*, 128(4), 224–233.
- [21] Kumar, P., Kumar, D., Jaipaul, A. and Tiwari, K.(2012). Evaporation Estimation Using Artificial Neural Networks and Adaptive Neuro-Fuzzy Inference System Techniques. *Journal of Meteorology*, 8(16), 88-97.
- [22] Mahdavi, M. (2003). *Applied Hydrology*. 5<sup>nd</sup> Ed., Tehran university.
- [23] Moghaddamnia, A., Gousheh, M. G., Piri, J., Amin, S. and Han, D. (2009). Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques. *Advances in Water Resources*, 32(1), 88-97
- [24] Mosaedi, A. and Ghobani sogh, M. (2010). Modification of RDI index Selecting the most appropriate distribution function In arid and semi-arid regions of Iran, *1st National Conference on Agrometeorology and Agricultural Water Management*, Karaj.
- [25] Nourani, V. and Sayyah Fard, M.(2012). Sensitivity analysis of the artificial neural network outputs in simulation of the evaporation process at different climatologic regimes. *Journal of Advances in Engineering Software* ,47(1), 127–146.
- [26] Piri, J. and Ansari, H. (2013). Daily pan evaporation modelling with ANFIS and NNARX. *Iran Agricultural Research*, 31(2), 51-64.
- [27] Piri, J., Mohammadi, K., Shamshirband, S. and Akib, S. (2016). Assessing the suitability of hybridizing the Cuckoo optimization algorithm with ANN and ANFIS techniques to predict daily evaporation. *Environmental Earth Sciences*, 75(3), 246.
- [28] Sayadi, H., Oulad ghfari,A., Faalian, A., and Sadroddini, A.A. (2010). Comparison of MLP and RBF neural networks for estimation of reference evapotranspiration. *Journal of Soil and Water*.19(1)1-12.
- [29] Serrano, M.L.A., Ruiz, A., Garcia, J.A., Anton, M. and Vaquero, J.M. (2005). Solar global radiation and sunshine duration in Extremadura (Spain). *Journal of Physica Scripta*,118,24-28.
- [30] Shafizade, F. and Mobin, M. H. (2009). Evaluation of drought in southern Iran Using the RDI index and Mann – Kendall, *Second National Conference on Water*, Behbahan.
- [31] Shayannejad, M. (2007). Comparative accuracy of artificial neural networks and Penman - mantis methods in calculating Potential evapotranspiration. *National Conference on Irrigation and Drainage Networks*. Shahid chamran university,ahvaz.Iran
- [32] Shiri ,J., Nazemi, A. H., Sadraddini, A., Landeras, G., Kisi, O., Fakheri Fard, A. and Marti, P.(2013). Global cross-station assessment of neuro-fuzzy models for estimating daily reference evapotranspiration. *Journal of Hydrology*, 480, 46–57
- [33] Singh,V. (1992). *Elementary Hydrology*, NJ, U.S.A.: Prentice Hall Inc.
- [34] Sudheer, K. P., Gosain, A. K., Mohana Rangan, D. and Saheb, S. M.(2002). Modelling evaporation using an artificial neural network algorithm. *Journal of Hydrological Processes*, 16, 3189–3202.

- [35] Takagi, T. and Sugeno, M. (1985). Fuzzy identification of systems and its application to modeling and control, *IEEE Transactions on Systems, Man, and Cybernetics*, 15, 116–132.
- [36] Trajkovic, S., Todorovic, B. and Stankovic, M. (2003). Forecasting of Reference Evapotranspiration by Artificial Neural Networks. *Journal of Irrigation and Drainage Engineering*, 129(6), 454–457.
- [37] Traore, S., Wang, Y.M. and Kerh, T. (2010). Artificial neural network for modeling reference evapotranspiration complex process in Sudano-Sahelian zone. *Journal of Agricultural Water Management*, 97 (5), 707–714.
- [38] Tsakiris, G. and Vangelis, H. (2005). Establishing a drought index incorporating evapotranspiration. *Journal of European water*, 9(10), 3-11.
- [39] Tsakiris, G., Pangalou, D. and Vangelis, H. (2006). Regional Drought Assessment Base on the Reconnaissance Drought Index (RDI). *Journal of Water Resource Management*, 21(5), 821-833.
- [40] Xu, C.Y. and Singh, V. P. (1998). Dependence of evaporation on meteorological variables at different time-scales and intercomparison of estimation methods. *Journal of Hydrol. Proc.*, 12(3), 429–442.
- [41] Zanetti, S.S., Sousa, E.F., Oliveira, V.P.S., Almeida, F.T. and Bernardo, S. (2007). Estimating evapotranspiration using artificial neural network and minimum climatological data. *Journal of Irrigation and Drainage Engineering*, 133 (2), 83–89.
- [42] Zare Abyaneh, H., Bayat Varkeshi, M., Marofi, S. and Amiri Chayjan, R. (2010). Evaluation of Artificial Neural Network and Adaptive Neuro Fuzzy Inference System in Decreasing of Reference Evapotranspiration Parameters. *Journal of Water and Soil*, 24(2), 297-305.