



# سومین همایش ملی دانش و فناوری علوم کشاورزی، منابع طبیعی و محیط زیست ایران

اهداء گواهینامه های معتبر بین المللی به شرکت کنندگان

www.mdconf.ir  
info@mdconf.ir

## مقایسه روش رگرسیون خطی چندگانه با روش های مبتنی بر هوش مصنوعی برای پیش بینی میزان مصرف آب شرب شهری، مطالعه موردی: شهرک مهدیه قم

مصطفی رضاعلی<sup>۱</sup>، عبدالرضا کریمی\*<sup>۲</sup>، بایرامعلی محمدنژاد<sup>۳</sup>، عبدالرضا رسولی کناری<sup>۴</sup>

۱- دانشجوی کارشناسی ارشد عمران- محیط زیست دانشگاه صنعتی قم، قم، ایران، Rezaali.m@qut.ac.ir

۲- استادیار گروه مهندسی عمران، دانشگاه صنعتی قم، قم، ایران، Karimi@qut.ac.ir

۳- استادیار گروه مهندسی عمران، دانشگاه صنعتی قم، قم، ایران، Mohammadnezhad@qut.ac.ir

۴- استادیار گروه مهندسی کامپیوتر، دانشگاه صنعتی قم، قم، ایران، Rasouli@qut.ac.ir

### چکیده

تنش ناشی از خشکسالی در ایران در ده های اخیر رو به افزایش بوده است. با افزایش جمعیت، افزایش میزان تقاضای آب و کاهش بارندگی، خشک سالی و کمبود آب بیشتر مورد توجه قرار می گیرد. یکی از راه های برون رفت از این بحران مدیریت عرضه و تقاضای آب است که لازم در کنار سایر اقدامات لازم برای مدیریت بحران آب مورد توجه قرار گیرد. بمنظور مدیریت عرضه و تقاضای آب لازم است تا میزان تقاضای آب آینده با دقت قابل قبولی پیش بینی شود. در دهه های اخیر، پژوهشگران با اولین کاربرد موفقیت آمیز هوش مصنوعی، به مقایسه و ارزیابی مدل های مبتنی بر هوش مصنوعی مانند شبکه عصبی مصنوعی با روش های آماری مبتنی بر رگرسیون خطی و غیرخطی پرداختند. در این پژوهش سعی شده با مقایسه عملکرد شبکه عصبی مصنوعی سری زمانی با مدل رگرسیون خطی چندگانه برای شهرک مهدیه قم، بهترین مدل موجود بررسی و انتخاب گردد. در نتیجه مدل شبکه عصبی مصنوعی سری زمانی با استفاده از الگوریتم های آموزشی لوبز-مارکوات، شبه نیوتنی و گرادیان کاهشی مورد بررسی قرار گرفت و سپس بهترین مدل توصیف کننده رفتار مصرف کننده انتخاب شد. مدل انتخاب شده سپس با مدل رگرسیون خطی چندگانه مورد مقایسه قرار گرفت و نقاط مثبت و منفی هر کدام تحلیل شد. نتایج نشان داد که مدل شبکه عصبی مصنوعی سری زمانی دقت بالاتری در مقایسه با مدل رگرسیون خطی چندگانه دارد.

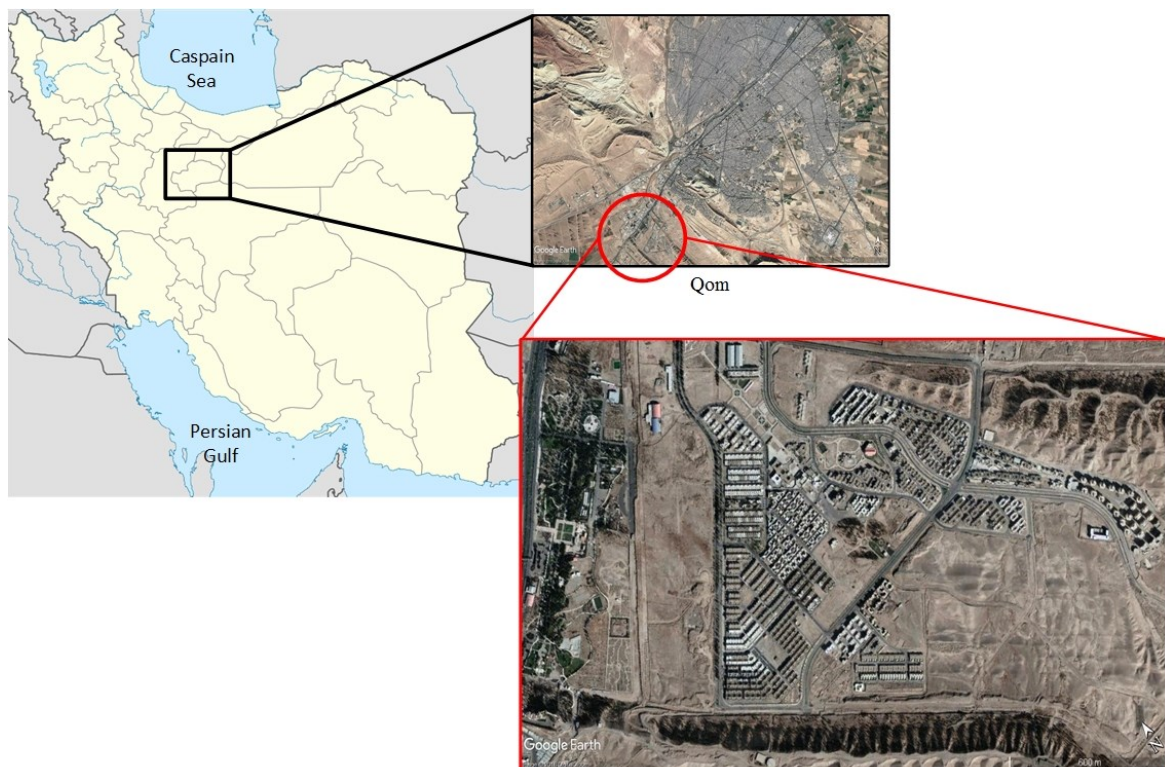
### واژه های کلیدی

پیش بینی تقاضا، شبکه عصبی مصنوعی، مدل رگرسیون خطی چندگانه، مدل های سری زمانی

### ۱- مقدمه و هدف

ایران هم اکنون در حال تجربه مشکلات جدی آب است. خشک سالی های مکرر توأم با برداشت بیش از حد آب های سطحی و زیرزمینی از طریق شبکه بزرگی از زیرساخت های هیدرولیکی و چاه های عمیق، وضعیت آب کشور را به سطح بحرانی رسانده است. از نشانه های این وضعیت خشک شدن دریاچه ها، رودخانه ها و تالاب ها، کاهش سطح

آب های زیرزمینی، فرونشست زمین، تخریب کیفیت آب، فرسایش خاک، بیابان زایی و طوفان های گردوغبار بیشتر است. یکی از استان هایی که در معرض بیشترین خطر خشک سالی است استان قم است. شهرک مهدیه قم با جمعیت بالغ بر سه هزار نفر شهرکی غالباً روحانی نشین به حساب می آید. اطلاعات میزان تقاضای آب این شهرک از این جهت مورد استفاده این پژوهش قرار گرفت که شهرک دارای شبکه توزیع آب جدا و زون بندی یکپارچه است. شکل ۱ تصویر محدوده تحت مطالعه را نشان می دهد.



شکل ۱: تصویر هوایی از شهرک مهدیه قم

با توجه به اینکه در اغلب نقاط شهر قم زون بندی شبکه به درستی انجام نشده است و اطلاعات فلو مترها ناقص است و یا اصلاً تجهیزات اندازه گیری دبی میزان تقاضا موجود نیست، به همین دلیل در این پژوهش منطقه تحت مطالعه به این شهرک محدود شد که هدف از این پژوهش برآورد میزان تقاضای آب این شهرک و مقایسه دقت مدل های شبکه عصبی مصنوعی سری زمانی با مدل رگرسیون چندگانه<sup>۱</sup> است. در پژوهش های مختلف شبکه عصبی مصنوعی بمنظور کاربردهای مختلفی مورد استفاده قرار گرفته است که از جمله آنها می توان به کنترل خودرو [۱]، پزشکی [۲]، امور تحصیلی [۳]، رباتیک [۴]، کارخانجات [۵]، کشاورزی [۶] و روان پزشکی [۷] تا استفاده از هوش مصنوعی در پیش بینی میزان تقاضای آب [۸]، پیش بینی دبی رودخانه ها [۹]، سطح آب زیرزمینی [۱۰] و سایر موارد اشاره کرد.

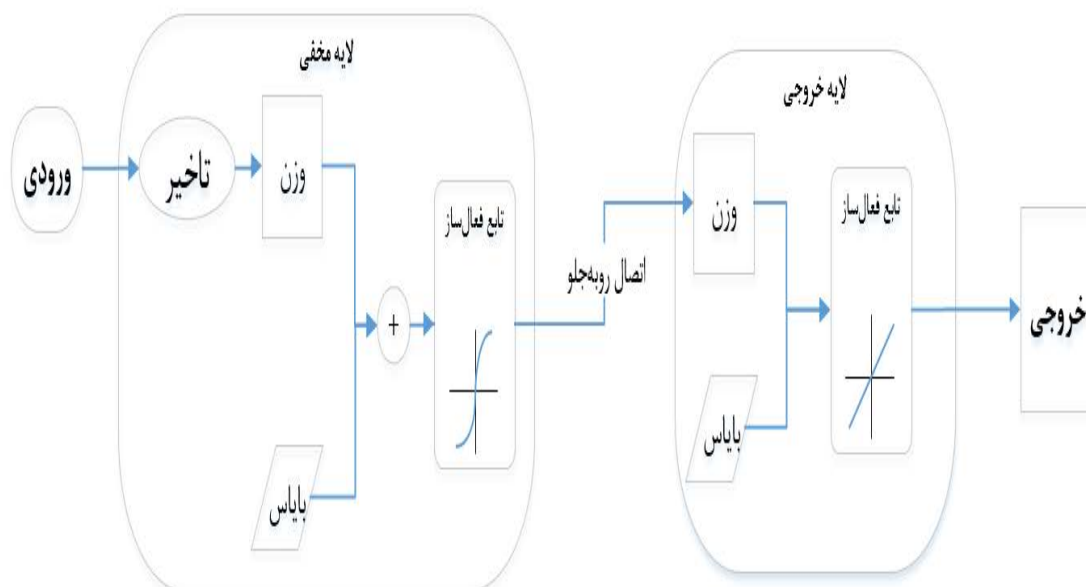
<sup>۱</sup> Multiple linear regression

## تئوری و پیشینه تحقیق

### شبکه عصبی مصنوعی

روش شبکه عصبی مصنوعی (ANN) اولین بار در سال ۱۹۴۳ میلادی توسط مک کلاچ و پیتز معرفی گردید و علی‌رغم به‌کارگیری یک ساختمان ساده از این مدل، سرعت و قدرت محاسباتی آن به‌شدت موردتوجه قرار گرفت. این روش تدریجاً در شاخه‌های متعدد علوم از جمله منابع آب کارایی خود را به اثبات رساند و مورد استفاده قرار گرفت.

یک شبکه عصبی مصنوعی، از سه لایه ورودی، خروجی و پردازش تشکیل می‌شود. هر لایه شامل گروهی از سلول‌های عصبی (نورون) است که عموماً با کلیه نورون‌های لایه‌های دیگر در ارتباط هستند، مگر این‌که کاربر ارتباط بین نورون‌ها را محدود کند. ولی نورون‌های هر لایه با سایر نورون‌های همان لایه، ارتباطی ندارند [۱۱]. نورون کوچک‌ترین واحد پردازشگر اطلاعات است که اساس عملکرد شبکه‌های عصبی را تشکیل می‌دهد. یک شبکه عصبی مجموعه‌ای از نورون‌ها است که با قرار گرفتن در لایه‌های مختلف، معماری خاصی را بر مبنای ارتباطات بین نورون‌ها در لایه‌های مختلف تشکیل می‌دهند. نورون می‌تواند یک تابع ریاضی غیرخطی باشد، در نتیجه یک شبکه عصبی که از اجتماع این نورون‌ها تشکیل می‌شود، نیز می‌تواند یک سامانه کاملاً پیچیده و غیرخطی باشد. در شبکه عصبی هر نورون به‌طور مستقل عمل می‌کند و رفتار کلی شبکه، برآیند رفتار نورون‌های متعدد است. به عبارت دیگر، نورون‌ها در یک روند همکاری، یکدیگر را تصحیح می‌کنند. شکل ۲ شمای کلی شبکه عصبی مصنوعی سری‌زمانی را نشان می‌دهد.



شکل ۲: شمای کلی شبکه عصبی مصنوعی پرسپترون چندلایه بکار رفته در پژوهش





## سومین همایش ملی دانش و فناوری علوم کشاورزی، منابع طبیعی و محیط زیست ایران

اهداء گواهینامه های معتبر بین المللی به شرکت کنندگان

www.mdconf.ir  
info@mdconf.ir

### روش های مبتنی بر رگرسیون

در بسیاری از مسائل هیدرولوژیکی مشاهده شده که یک متغیر مشخص مثل  $y$  به یک یا چند متغیر مستقل مثل  $x_1, x_2, \dots, x_n$  وابسته است و می توان با داشتن مقادیر آن ها، مقدار  $y$  را تخمین زد. روشی که برای تعیین رابطه بین متغیرها در این گونه موارد مورد استفاده قرار می گیرد، روش رگرسیون (تحلیل رگرسیون) یا تحلیل همبستگی (*Regression Analyss*) نامیده می شود [۱۲].

چنانچه مقادیر  $x$  و  $y$  بر روی یک صفحه در مقابل یکدیگر ترسیم شوند، یک نمودار پراکندگی به وجود می آید که در حالت خاص ممکن است نمودار پراکندگی حاصل، دلالت بر وجود رابطه ای خطی بین متغیرها به صورت فرمول ۱:

$$y = a + bx \quad (1)$$

و یا به صورت فرمول ۲:

$$y = a + b_1x_1 + b_2x_2 + \dots + b_nx_n \quad (2)$$

داشته با شد. برای تحلیل همبستگی این داده ها باید بهترین خط (یا منحنی) را که می توان بر داده ها برازش داد تعیین نمود. هم چنین باید معیاری برای سنجش میزان پراکندگی نقاط در اطراف خط برازش داده شده یا به عبارت دیگر دقت نتایجی که با استفاده از خط رگرسیون به دست می آید تعیین شود. به منظور تعیین بهترین خط (یا منحنی) معمولاً از روش حداقل مربعات (*Least Squares Method*) استفاده می شود. در این روش یک خط (یا منحنی) طوری بر داده ها برازش داده می شود که مجموع مربعات انحراف از کلیه نقاط خط رگرسیون به حداقل خود برسد.

### مواد و روش ها

داده های سری زمانی میزان مصرف آب شرب مخزن مهدیه که هر ۱۵ دقیقه یکبار به مدت ۲۸۳ روز به ثبت رسیده از سازمان آب و فاضلاب شهر قم دریافت شد. بمنظور محاسبه دقیق میزان مصرف آب، از داده های ۱۵ دقیقه ای به صورت روزانه انتگرال گرفته شد. این روش نسبت به روش میانگین روزانه از دقت بالاتری برخوردار است. به دلیل بالا بودن تعداد روزها، در محیط *MATLAB* برنامه ای نوشته شد تا این عملیات را به صورت خودکار برای هر روز انجام دهد. داده های مذکور به همراه داده های جوی شامل حداکثر دمای روزانه، سرعت باد، میزان بارندگی، وقوع یا عدم وقوع بارندگی و تعطیلات و مناسبت های تقویم، تحت آنالیز *Spearman-correlation* قرار گرفت، تا مشخص شود کدام یک از داده ها سطح معناداری بیشتری نسبت به سایر پارامترها دارند. نتایج نشان داد که پارامترهای میزان بارندگی، وقوع یا عدم وقوع بارندگی و مناسبت ها به ترتیب دارای کمترین میزان معناداری هستند. در نتیجه از این پارامترها به عنوان ورودی مدل ها استفاده نشد. در مرحله بعد به منظور شناخت بهترین تأخیرها برای اعمال روی داده های سری زمانی، آنالیز *Cross-Correlation* بر روی داده ها انجام شد. در نهایت از نتایج این آنالیز به منظور اعمال تأخیرها بر روی مدل های سری زمانی استفاده شد.



تمام داده‌های ورودی به مدل عصبی به عنوان ورودی مدل رگرسیون چندگانه با فاصله اطمینان ۹۵ درصد استفاده شد. همچنین داده‌های ۳۰ روز بصورت تصادفی انتخاب شد تا مدل رگرسیونی بوسیله آن اعتبارسنجی شود.

### نتایج و بحث

تعداد نرون‌های یک مدل شبکه عصبی مصنوعی به شدت وابسته به رفتار داده‌ها و تعداد داده‌های تحت آموزش است. این بدین معنی است که هرچه رفتار یک تابع غیرخطی تر یا به تعبیری رندم‌تر باشد، مدل در تقریب آن تابع بیشتر نیاز به داده‌های آموزش دارد و در نتیجه تعداد نرون‌های لازم برای آن تابع نیز بیشتر است. برای تحقیق کنونی بمنظور پیدا کردن بهترین توپولوژی از مدل شبکه عصبی مصنوعی دینامیکی، مدل برای هر شبیه‌سازی با تعداد نرون مختلف آموزش داده شد. همچنین بمنظور کاهش هزینه‌های محاسباتی حداکثر تعداد نرون‌ها با توجه به تعداد داده‌های ورودی ۱۵ عدد قرار داده شد. جدول ۱ میزان خطاها و ضریب رگرسیون هر مدل با الگوریتم‌های آموزش لونبرگ-مارکوارت ( $LM$ )، کاهش گرادیان پس‌انتشارخطا ( $GBP$ ) و الگوریتم شبه‌نیوتنی ( $BFG$ ) را نشان می‌دهد.

جدول ۱: نتایج اجراهای پیاپی از مدل شبکه عصبی مصنوعی سری‌زمانی با توپولوژی‌های مختلف

توپولوژی شبکه	الگوریتم	ضریب همبستگی	RMSE آموزش	RMSE اعتبارسنجی	RMSE تست
4_1_1	LM	0.799	3.91E-04	2.53E-04	3.36E-04
4_2_1	LM	0.829	3.41E-04	2.54E-04	3.62E-04
4_3_1	LM	0.834	3.32E-04	2.61E-04	3.40E-04
4_4_1	LM	0.746	5.01E-04	2.65E-04	3.46E-04
4_5_1	LM	0.856	2.87E-04	2.75E-04	3.76E-04
4_6_1	LM	0.171	6.63E-04	3.01E-04	4.62E-04
4_7_1	LM	0.773	3.82E-04	3.32E-04	3.80E-04
4_8_1	LM	-0.651	1.68E-03	5.91E-04	5.15E-04
4_9_1	LM	0.574	7.04E-04	6.30E-04	7.73E-04
4_10_1	LM	-0.630	1.83E-03	6.92E-04	5.56E-04
4_11_1	LM	-0.520	1.61E-03	8.90E-04	7.05E-04
4_12_1	LM	0.213	1.65E-03	1.26E-03	1.10E-03
4_13_1	LM	0.724	5.66E-04	2.51E-04	3.87E-04
4_14_1	LM	0.813	4.13E-04	3.39E-04	3.76E-04

4_15_1	LM	0.757	3.90E-04	3.32E-04	4.18E-04
4_1_1	GBP	0.64	4.87E-04	3.35E-04	3.97E-04
4_2_1	GBP	-0.21	2.11E-03	1.53E-03	1.38E-03
4_3_1	GBP	0.58	1.37E-03	1.13E-03	1.15E-03
4_4_1	GBP	0.65	8.85E-04	8.65E-04	7.40E-04
4_5_1	GBP	0.53	5.70E-04	3.17E-04	3.58E-04
4_6_1	GBP	-0.56	9.45E-04	1.08E-03	1.28E-03
4_7_1	GBP	0.51	9.37E-04	7.82E-04	7.32E-04
4_8_1	GBP	0.02	6.95E-04	4.76E-04	6.17E-04
4_9_1	GBP	-0.16	1.03E-03	4.10E-04	4.63E-04
4_10_1	GBP	0.66	5.77E-04	5.06E-04	5.23E-04
4_11_1	GBP	-0.39	8.55E-04	1.19E-03	1.32E-03
4_12_1	GBP	-0.06	8.34E-04	7.49E-04	8.20E-04
4_13_1	GBP	0.08	6.63E-04	4.84E-04	6.28E-04
4_14_1	GBP	-0.53	1.26E-03	4.69E-04	3.74E-04
4_15_1	GBP	-0.25	1.88E-03	1.24E-03	1.10E-03
4_1_1	BFG	-0.46	1.56E-03	7.04E-04	4.52E-04
4_2_1	BFG	0.74	4.24E-04	2.53E-04	3.51E-04
4_3_1	BFG	0.80	3.54E-04	2.69E-04	4.09E-04
4_4_1	BFG	0.77	3.87E-04	2.81E-04	4.03E-04
4_5_1	BFG	0.80	3.73E-04	2.55E-04	3.53E-04
4_6_1	BFG	-0.25	8.17E-04	4.44E-04	6.18E-04
4_7_1	BFG	0.81	3.63E-04	2.53E-04	3.27E-04
4_8_1	BFG	0.74	4.06E-04	2.93E-04	3.90E-04
4_9_1	BFG	0.78	3.80E-04	2.48E-04	3.91E-04
4_10_1	BFG	0.75	4.09E-04	2.93E-04	3.49E-04
4_11_1	BFG	0.69	4.56E-04	2.68E-04	3.61E-04
4_12_1	BFG	0.72	4.06E-04	4.04E-04	3.96E-04
4_13_1	BFG	0.69	4.55E-04	2.86E-04	3.74E-04
4_14_1	BFG	0.69	4.41E-04	3.18E-04	4.24E-04
4_15_1	BFG	-0.53	1.49E-03	7.24E-04	5.51E-04

با مقایسه نتایج بدست آمده از هر مدل مشخص می شود که مدل شبکه عصبی مصنوعی با الگوریتم یادگیری لونیگ-مارکوات عملکرد بهتری نسبت به سایر مدل ها داشته است. همچنین مشخص می شود که مدل آموزش دیده با این الگوریتم با ۵ عدد نوروں توانایی تقریب بهتری نسبت به سایر مدل ها دارد. همچنین نتایج مدل رگرسیون خطی چندگانه در مقایسه با مدل شبکه عصبی مصنوعی سری زمانی آورده شده است.

جدول ۲: مقایسه نتایج بهترین مدل شبکه عصبی مصنوعی سری زمانی با مدل رگرسیون چندگانه خطی

مدل	RMSE آموزش	RMSE اعتبارسنجی
رگرسیون چندگانه خطی	2.75E-04	2.87E-04
شبکه عصبی مصنوعی سری زمانی	4.25E-01	4.25E-01

همانطور که جدول ۲ نشان می دهد مدل شبکه عصبی مصنوعی سری زمانی عملکرد به مراتب بهتری نسبت به مدل رگرسیون چندگانه خطی دارد. یکی از مهم ترین عوامل این برتری را می توان در آموزش پیاپی این مدل دانست که می تواند در هر دوره یا *Epoch* عملکرد خود را باتوجه به تابع هزینه بهبود بخشد.

## نتیجه گیری و جمع بندی

پیش بینی نیاز آب شهری یکی از ملزومات کنترل و مدیریت عرضه و تقاضا در شبکه توزیع آب شهری است. بدین منظور مدل های مبتنی بر روش های آماری و اخیراً مدل های مبتنی بر هوش مصنوعی تحت پژوهش های گسترده قرار گرفته است. در این پژوهش دو نوع از پرکاربردترین مدل ها که عبارت اند از مدل رگرسیون چندگانه خطی و مدل شبکه عصبی مصنوعی سری زمانی در زمینه پیش بینی تقاضا در شبکه آب شهری تحت بررسی قرار گرفت تا مشخص شود کدام یک عملکرد بهتری از نظر میزان خطای پیش بینی دارند. نتایج نشان داد که مدل شبکه عصبی مصنوعی سری زمانی برتری قابل توجهی نسبت به مدل رگرسیون خطی چندگانه دارد که این برتری می تواند دلایل متعددی داشته باشد. از مهم ترین این دلایل را می توان ماهیت غیرخطی و پروسه آموزش مدل های عصبی دانست.

## پیشهادات

سایر پژوهش ها می تواند دربرگیرنده استفاده از سایر مدل های آماری نظیر *ARIMA* یا *SARIMA* و مقایسه آن با شبکه عصبی مصنوعی با ترکیب توابع انتقال مختلف باشد.



## سومین همایش ملی دانش و فناوری علوم کشاورزی، منابع طبیعی و محیط زیست ایران

اهداء گواهینامه های معتبر بین المللی به شرکت کنندگان

www.mdconf.ir  
info@mdconf.ir

### تشکر و قدردانی

نویسندگان مقاله از شرکت آب و فاضلاب شهری استان قم و سازمان هواشناسی کشوری بمنظور مساعدت در دریافت داده ها تشکر می نمایند.

### مراجع

- [1] S. Tsugawa, T. Yatabe, T. Hirose, and S. Matsumoto, "An automobile with artificial intelligence," in *Proceedings of the 6th international joint conference on Artificial intelligence-Volume 2*, 1979, pp. 893-895: Morgan Kaufmann Publishers Inc.
- [2] P. Szolovits, *Artificial intelligence in medicine*. Westview Press Boulder, CO, 1982.
- [3] S. P. Lajoie and M. Vivet, *Artificial Intelligence in Education*. IOS Press, 2002.
- [4] M. Brady, "Artificial intelligence and robotics," *Artificial intelligence*, vol. 26, no. 1, pp. 79-121, 1985.
- [5] W. I. Bullers, S. Y. Nof, and A. B. Whinston, "Artificial intelligence in manufacturing planning and control," *Aiie transactions*, vol. 12, no. 4, pp. 351-363, 1980.
- [6] M. Aitkenhead, I. Dalgetty, C. Mullins, A. McDonald, and N. Strachan, "Weed and crop discrimination using image analysis and artificial intelligence methods," *Computers and electronics in Agriculture*, vol. 39, no. 3, pp. 157-171, 2003.
- [7] M. A. Boden, "Artificial intelligence in psychology: interdisciplinary essays," 1989.
- [8] A. Jain and L. E. Ormsbee, "Short-term water demand forecast modeling techniques—Conventional methods versus AI," *Journal (American Water Works Association)*, vol. 94, no. 7, pp. 64-72, 2002.
- [9] C. Imrie, S. Durucan, and A. Korre, "River flow prediction using artificial neural networks: generalisation beyond the calibration range," *Journal of hydrology*, vol. 233, no. 1-4, pp. 138-153, 2000.
- [10] I. N. Daliakopoulos, P. Coulibaly, and I. K. Tsanis, "Groundwater level forecasting using artificial neural networks," *Journal of Hydrology*, vol. 309, no. 1-4, pp. 229-240, 2005.
- [11] H. R. Maier and G. C. Dandy, "Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications," *Environmental Modelling & Software*, vol. 15, no. 1, pp. 101-124, 2000/01/01/ 2000.
- [12] S. Ho, M. Xie, and T. Goh, "A comparative study of neural network and Box-Jenkins ARIMA modeling in time series prediction," *Computers & Industrial Engineering*, vol. 42, no. 2-4, pp. 371-375, 2002.