

Comparison of Artificial Neural Network and Multiple Linear Regressions in Predicting the Turbidity of Slow Sand Filtration of Tabas Water Treatment Plant “ANN and MLR Application in Turbidity Prediction”

Mohsen Niazi

MS.c Student, Environmental Health Engineering, Birjand University of Medical Sciences, birjand, Iran.

Ali Naghizadeh

Associate Professor, Medical Toxicology and Drug Abuse Research Center (MTDRC), Birjand University of Medical Sciences (BUMS), Birjand, Iran.

Mansour Baziar

Assistant Professor Department of Environmental Health Engineering, Ferdows School of Allied Medicine and Public Health, Birjand University of Medical Sciences, Birjand, Iran. (Corresponding Author):
Email: baziar.ehe@gmail.com

Received: 2021/10/26

Accepted: 2022/01/18

Document Type: Research article

ABSTRACT

Background and Aim: The turbidity of treated water is measured as an important parameter to determine the quality of drinking or industrial water in all treatment plants. In research on the relevance of pathogens like Giardia and Cryptosporidium, which cause hazardous illnesses like dysentery, the link between lowering turbidity and improving the clearance of these bacteria has been shown.

Materials and Methods: An artificial neural network (ANN) model and multiple linear regression (MLR) were developed, and their performance was compared to predict the turbidity of treated water at Tabas water treatment plant. Total dissolved solids, pH, temperature and input turbidity of raw water were used as the input parameters of models in the predictions. The best backpropagation algorithm and number of neurons were determined to optimize the model architecture.

Results: The results showed that the Levenberg–Marquardt algorithm was selected as the best algorithm and the number of optimal neurons was determined to be 16.

The sensitivity analysis of the neural network model also revealed that the input turbidity, with a value of 29%, is the most essential parameter in the creation of the ANN model.

Conclusion: The results of the correlation coefficient of MLR and ANN models were obtained to obtain data 0.63 and 0.8921 and for testing data 0.60 and 0.8571, respectively, which show the superiority of ANN model to predict the turbidity of the output of Tabas water treatment plant.

Keywords: Turbidity, water treatment, artificial neural network, multiple linear regression, prediction

► **Citation:** Niazi M, Naghizadeh A, Baziar M. Comparison of artificial neural network and multiple linear regressions in predicting the turbidity of slow sand filtration of Tabas water treatment plant “ANN and MLR application in turbidity prediction”. *Iranian Journal of Research in Environmental Health*. Spring 2022; 8(1): 33-45.

مقایسه شبکه عصبی مصنوعی و رگرسیون خطی چندگانه در پیش‌بینی کدورت فیلتراسیون شنی کند تصفیه‌خانه آب طبرس "کاربرد ANN و MLR در پیش‌بینی کدورت"

چکیده

زمینه و هدف: کدورت آب تصفیه شده به‌عنوان یک پارامتر مهم در تعیین کیفیت آب آشامیدنی و یا صنعتی در تمامی تصفیه‌خانه‌ها اندازه‌گیری می‌شود. از دبرباز با توجه به اهمیت یافتن شیوع عوامل بیماری‌زا مانند زیاردیا و کریپتوسپوریدیوم که عامل ایجاد بیماری‌های خطرناکی همچون اسهال خونی می‌باشند، رابطه کاهش میزان کدورت و افزایش حذف این میکروارگانیسم‌ها در مطالعات به اثبات رسیده است.

مواد و روش‌ها: در این مطالعه یک مدل شبکه عصبی مصنوعی و رگرسیون خطی چندگانه در پیش‌بینی کدورت خروجی از آب تصفیه شده تصفیه‌خانه شهر طبرس توسعه و عملکرد آنها با هم مقایسه گردید. کل جامدات محلول، pH، دما و کدورت ورودی به‌عنوان پارامترهای ورودی مدل‌ها در پیش‌بینی‌ها استفاده شد. بهترین الگوریتم پس‌انتشار و تعداد نورون برای بهینه‌سازی معماری مدل تعیین شد. **یافته‌ها:** بر اساس نتایج، الگوریتم لوبنبرگ-مارکوارت به‌عنوان بهترین الگوریتم انتخاب شد و تعداد نورون بهینه نیز ۱۶ تعیین شد. همچنین نتایج تحلیل حساسیت مدل شبکه عصبی نشان داد که کدورت ورودی با مقدار ۲۹٪ به‌عنوان مهم‌ترین پارامتر در توسعه مدل شبکه عصبی مصنوعی است. **نتیجه‌گیری:** نتایج ضریب همبستگی مدل رگرسیون خطی چندگانه و مدل شبکه عصبی مصنوعی برای داده‌های آموزش به‌ترتیب ۰/۶۳ و ۰/۸۹۲ و برای داده‌های تست ۰/۶۰ و ۰/۸۵۷۱ به‌دست آمد که نشان از برتری مدل شبکه عصبی مصنوعی در پیش‌بینی کدورت خروجی از تصفیه‌خانه آب طبرس داشت.

کلید واژه‌ها: پیش‌بینی، تصفیه‌خانه آب، رگرسیون خطی چندگانه، شبکه عصبی مصنوعی، کدورت

محسن نیازی

دانشجوی کارشناسی ارشد، مهندسی بهداشت محیط، دانشکده بهداشت، دانشگاه علوم پزشکی بیرجند، بیرجند، ایران.

علی نقی‌زاده

دانشیار مهندسی بهداشت محیط، مرکز تحقیقات سم‌شناسی پزشکی و سوء مصرف مواد مخدر (MTDRC)، دانشکده بهداشت، دانشگاه علوم پزشکی بیرجند، بیرجند، ایران.

منصور بازاریار

* استادیار مهندسی بهداشت محیط، مرکز تحقیقات سم‌شناسی پزشکی و سوء مصرف مواد مخدر (MTDRC)، دانشکده بهداشت، دانشگاه علوم پزشکی بیرجند، بیرجند، ایران.

استادیار مهندسی بهداشت محیط، دانشکده پیراپزشکی و بهداشت فردوس، دانشگاه علوم پزشکی بیرجند، بیرجند، ایران. (نویسنده مسئول): پست الکترونیک: baziar.ehe@gmail.com

تاریخ دریافت: ۱۴۰۰/۰۸/۱۳

تاریخ پذیرش: ۱۴۰۰/۱۰/۲۸

نوع مقاله: مقاله پژوهشی

◀ **استناد:** نیازی م، نقی‌زاده ع، بازاریار م. مقایسه شبکه عصبی مصنوعی و رگرسیون خطی چندگانه در پیش‌بینی کدورت فیلتراسیون شنی کند تصفیه‌خانه آب طبرس "کاربرد ANN و MLR در پیش‌بینی کدورت". فصلنامه پژوهش در بهداشت محیط. بهار ۱۴۰۱؛ ۸(۱): ۳۳-۴۵.

مقدمه

طبق تعریف استاندارد، کدورت به عنوان یک خاصیت فیزیکی نمونه آبی است که باعث می شود درصدی از نور تابیده شده به جای انتقال در خطوط مستقیم از طریق نمونه، پراکنده و جذب شود (۱). عموماً کدورت آب توسط مواد معلق همچون رس، سیلت، رسوبات شیمیایی آهن و منگنز، مواد آلی یا معدنی، بقایای گیاهان و میکروارگانیسم‌ها ایجاد می شود (۲). کدورت آب تصفیه شده به عنوان یک پارامتر مهم در تعیین کیفیت آب آشامیدنی و یا صنعتی در تمامی تصفیه خانه‌ها اندازه گیری می شود. کدورت به خودی خود خطر خاصی را برای سلامت انسان ایجاد نمی کند، ولی می تواند نمایانگر کیفیت پایین آب باشد. از دیرباز با توجه به اهمیت یافتن شیوع عوامل بیماری‌زا مانند ژیاوریا و کریپتوسپوریوم که عامل ایجاد بیماری‌های خطرناکی همچون اسهال خونی می باشند، رابطه کاهش میزان کدورت و افزایش حذف این میکروارگانیسم‌ها در مطالعات به اثبات رسیده است، بنابراین سازمان بهداشت جهانی (WHO) ^۱ توصیه می کند که کدورت آب نباید از سطح ۱ واحد کدورت نفلومتری (NTU) ^۲ قبل از کلرزنی تجاوز نماید (۳).

معمولاً مقادیر کدورت آب در طی زمان به آرامی تغییر می کند که البته این تغییر به خصوصیات و تغییرات حوضچه‌های آبریز بستگی دارد. مقادیر کدورت، گاهی اوقات در یک دوره کوتاه و به صورت تصادفی ممکن است به بالاترین مقدار خود برسد که این مقادیر بالا با رویدادهای محیطی مانند بارندگی شدید مرتبط است، اما می تواند نتیجه فعالیت‌های عملیاتی مانند پمپاژ نیز باشد. پیک در کدورت، یک چالش مهم برای عملکرد صحیح یک شرکت تأمین آب آشامیدنی است (۴).

کدورت آب از چند جنبه می تواند مهم باشد: الف) شفافیت آب: به طور کلی زیبایی آب یعنی صاف و زلال بودن آن است که این ویژگی با عدم وجود کدورت آب رابطه مستقیم دارد. از دیدگاه

مصرف کننده، صاف و شفاف بودن آب مترادف با بهداشتی بودن آن است. ب) قابلیت صاف کردن آب: کدورت ناشی از ذرات معلق، قابلیت صاف شدن دارند، در حالی که کدورت ناشی از ذرات کلوئیدی می توانند مشکلات فراوانی را در حذف ایجاد نمایند. ج) گندزدایی: کاملاً مشهود است که عمل گندزدایی آب در آب‌های با کدورت بالا به دلیل چسبیدن میکروارگانیسم‌ها به عوامل ایجاد کننده کدورت به طور مؤثری انجام نمی گیرد که دلیل آن، عدم تماس مستقیم میکروارگانیسم‌ها با این مواد گندزدا است. از این رو زدودن کدورت آب کمک شایانی به بهبود گندزدایی آب می کند (۵).

سیستم فیلتر شنی کند، یکی از اولین فرآیندهایی است که برای حذف آلاینده‌ها از آب‌های سطحی برای تولید آب آشامیدنی استفاده شده است. فیلترهای شنی کند به دلیل سادگی، کارایی و اقتصادی بودن، ابزار مناسبی برای تصفیه آب، به ویژه تأمین آب جوامع در حال توسعه است. فیلترهای شنی کند در یک نرخ فیلتراسیون بسیار پایین (۰/۱ میلی لیتر در ساعت) بهره‌برداری می شوند. معمولاً در این سیستم‌ها، عمل پیش کلرزنی انجام نمی شود و اندازه ماسه مورد استفاده آنها نیز ۰/۲ میلی متر است. اساساً، تصفیه آب در چنین فیلترهایی، نتیجه عبور آب از میان فیلتر توسعه داده شده در چند میلی متر بالای واحد فیلتراسیون همراه با فعالیت بیولوژیکی است. بنابراین، هر دو مکانیسم فیزیکی و بیولوژیکی در جذب ذرات در فیلتراسیون کند مهم هستند (۶، ۷).

این نوع فیلترها به صورت کاملاً مؤثر، باکتری‌ها، ویروس‌ها، پروتوزا، کدورت و فلزات سنگین را حذف می کنند، اما فضای زیادی را اشغال می کنند. از طرفی عملکرد آنها به شدت تحت تأثیر دمای محیط است. اگر دمای محیط کاهش قابل ملاحظه‌ای داشته باشد، به علت آسیب به لایه بیولوژیکی، بازده فیلترهای شنی کند کاهش می یابد (۸). همچنین ورود آب سخت، آب با کدورت غیرمعمول و همچنین آب با میزان جامدات محلول بالا مانند سدیم، نیترات، نیتريت، سولفات و فلوراید، از عملکرد صحیح این

1. World Health Organization
2. Nephelometric Turbidity unit

نوع از فیلترها می‌کاهد. به علت نیاز روزافزون به آبی که ویژگی‌های فیزیکی، شیمیایی و بیولوژیکی مناسب شرب را داشته باشد، تأمین چنین آبی به عملکرد صحیح تصفیه‌خانه آب وابسته است (۶). بنابراین بررسی رفتار عملکردی تصفیه‌خانه بر اساس معیار کیفی آب با استفاده از مدل‌های تخمین‌گرا می‌تواند مفید باشد. از طرفی، تصفیه آب در تصفیه‌خانه شامل چندین مرحله فیزیکی و شیمیایی و بیولوژیکی می‌باشد که معمولاً از طریق انجام آزمایشات و یا توسط اپراتور انجام می‌شود که به دلیل مداخلات عوامل مختلف در فرآیند تصفیه، کنترل این اعمال و پیش‌بینی رفتار سیستم تصفیه‌خانه بسیار مشکل است. از این رو طراحی مدلی برای پیش‌بینی میزان کدورت خروجی فیلتراسیون، کمک شایانی به بهبود فرآیند تصفیه می‌کند.

محققان زیادی استفاده از روش‌های هوش مصنوعی مانند شبکه‌های عصبی مصنوعی را در مباحث آلودگی محیط زیست به واسطه پیچیدگی سیستم‌های طبیعی و قابل اعتماد بودن این روش‌ها توصیه کرده‌اند. شبکه‌های عصبی مصنوعی به‌عنوان یکی از روش‌های محاسباتی در هوش مصنوعی، یک سیستم و تکنیک جدید محاسباتی هستند که از سیستم عصبی انسان الهام گرفته و داده‌ها یا اطلاعات را پردازش می‌کنند (۹، ۱۰). بخش کلیدی این سیستم محاسباتی، پردازنده‌های اطلاعاتی آنها است که نورون نامیده می‌شوند. سیستم شبکه عصبی مصنوعی (ANN) ^۱ از تعداد زیادی نورون به هم پیوسته تشکیل شده است که برای حل یک مشکل با هم اعمال نقش می‌کنند (۱۱). شبکه‌های عصبی مصنوعی به دلیل کاربر پسند بودن در شبیه‌سازی، عملکرد پیش‌بینی بالاتر در مدل‌سازی و پیش‌بینی بهتر نسبت به روش‌های استاندارد، به‌عنوان ابزارهای قابل اعتماد در علوم مختلف شناخته می‌شوند. شبکه‌های عصبی مصنوعی برای پیش‌بینی چندین شاخص کیفیت آب مانند سولفات، نیترات، نسبت جذب سدیم (SAR) ^۲، پیش‌بینی غلظت آهن در رودخانه منتج به تصفیه‌خانه آب شهر بغداد (عراق)، پیش‌بینی عملکرد تصفیه‌خانه

فاضلاب و غیره استفاده شدند (۱۲). با این حال، هیچ مطالعه‌ای در مورد کاربرد ANN در پیش‌بینی عملکرد فیلترهای شنی در تصفیه‌خانه آب وجود ندارد. از طرفی رگرسیون خطی چندگانه ^۳ نیز به‌عنوان یک روش استاندارد برای مقایسه عملکرد در شبیه‌سازی انتخاب گردید. بنابراین مطالعه حاضر با هدف مقایسه مدل‌هایی بر پایه شبکه عصبی مصنوعی و رگرسیون خطی چندگانه با استفاده از پارامترهای کیفی آب ورودی به تصفیه‌خانه (pH، دما، کل جامدات محلول و کدورت ورودی) و مقایسه توسعه و عملکرد آنها در پیش‌بینی کدورت خروجی از فیلتراسیون شنی و همچنین انجام یک تحلیل حساسیت برای تعیین درصد اهمیت هر متغیر در پیش‌بینی کدورت خروجی برای مدل برتر انجام شد.

روش کار

مشخصات تصفیه‌خانه طبس

تصفیه‌خانه آب طبس جهت تصفیه آب سد نهرین در مختصات با طول جغرافیایی ۵۶ درجه و ۹۹ دقیقه شرقی و عرض جغرافیایی ۳۳ درجه و ۶۱ دقیقه شمالی در شرق شهر طبس قرار دارد. آب ورودی به تصفیه‌خانه پس از طی مراحل انعقاد، لخته‌سازی، فیلتراسیون و گندزدایی نهایی از طریق خط لوله‌ای به طول حدود ۵ کیلومتر وارد مخزن ذخیره شده و آب مصرفی حدود ۵۵۰۰۰ نفر را تأمین می‌کند. نوع زلال‌ساز تصفیه‌خانه پولساتور می‌باشد. شماتیک مراحل مختلف تصفیه آب در تصفیه‌خانه آب شهر طبس در شکل ۱ نشان داده شده است. سنجش کل جامدات محلول توسط TDS Meter پرتابل لابلستر ^۴ و سنجش pH، دما و کدورت توسط دستگاه آنالاین اندرس هاورز ^۵ انجام شد.

شبکه عصبی مصنوعی

در این تحقیق، یک مدل ANN پیش‌خور برای پیش‌بینی کدورت خروجی فیلتراسیون شنی تصفیه‌خانه آب طبس (ایران) توسعه داده شد (شکل ۲). در طراحی ساختار و معماری مدل، تعداد

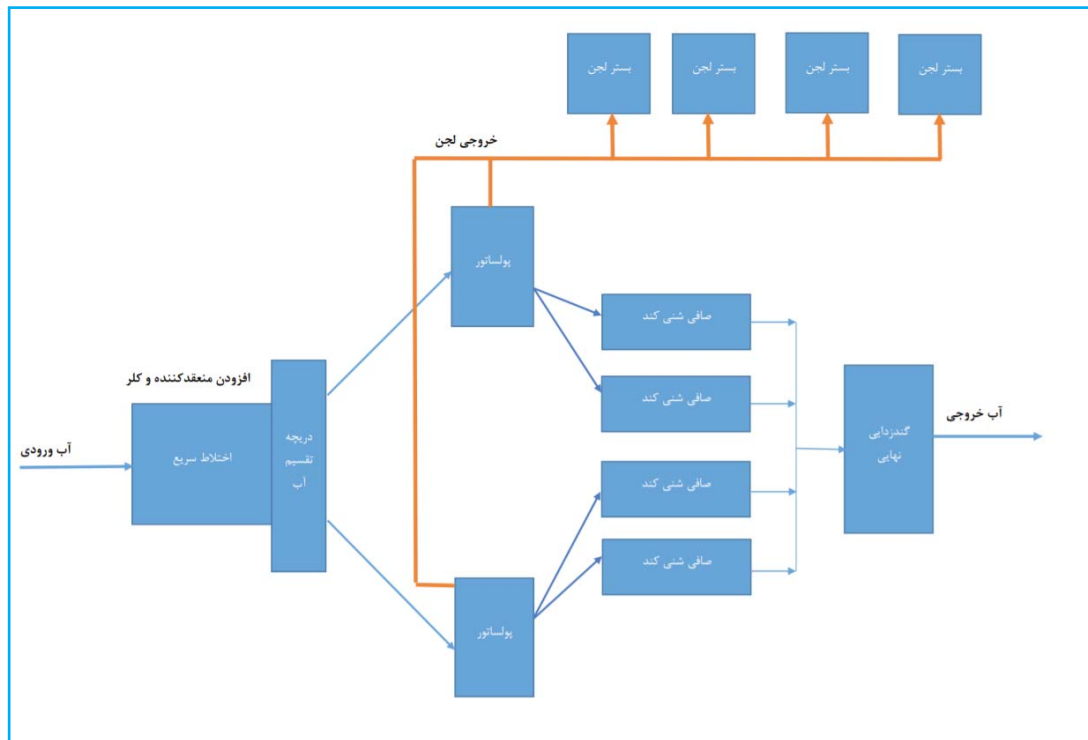
3. Multiple linear regression

4. Lobster

5. Endress+Hauser

1. Artificial neural network

2. Sodium adsorption ratio



شکل ۱. شماتیک مراحل مختلف تصفیه آب در تصفیه‌خانه آب شهر تبریز

gradient (trainscg) scaled - resilient backpropagation (trainrp) - Polak-Ribière conjugate gradient (traincgp) - one step secant (trainoss) - BFGS quasi-Newton (trainbfg) - gradient descent with momentum (traingdm)

با تعداد و مقادیر یکسان داده‌ها استفاده شد. همچنین تابع انتقال سیگموئید (Tansig)^۲ و تابع انتقال خطی (purelin)^۳ به ترتیب برای لایه‌های پنهان و خروجی مدل‌های توسعه داده شده اتخاذ گردید. مقادیر MSE با استفاده از معادله MSE تعیین شد. برای شناسایی مناسب‌ترین تعداد نورون برای لایه پنهان مدل شبکه عصبی مصنوعی، مقادیر ۱ تا ۲۰ نورون مورد آزمایش قرار گرفت. هدف از این مطالعه، به حداقل رساندن خطا و افزایش دقت پیش‌بینی خروجی مدل توسعه داده شده بود. بنابراین، این روش مدل‌سازی با ۱۰ تکرار در ۳ مرحله اعتبارسنجی، آزمایش و آموزش انجام گردید.

نورون‌ها در لایه‌های خروجی و ورودی مطابق با تعداد متغیرهای خروجی و ورودی مدل در نظر گرفته شد. بر این اساس، تعداد زیادی از مدل‌های شبکه عصبی مصنوعی با اعمال نورون‌های متفاوت در لایه پنهان شبکه عصبی در فرآیند آموزش مدل ساخته شد. در مراحل اعتبارسنجی و آزمون مدل، دقت مدل‌ها بر اساس شاخص‌های ارزیابی مانند میانگین مربعات خطا (MSE)^۱ و ضریب رگرسیون (R) مدل بررسی شد. در فرآیند مدل‌سازی، ابتدا کل داده‌ها به نسبت ۸۰:۲۰ به دو بخش تقسیم شدند. سپس بخش اول داده‌ها (۸۰٪) به سه بخش تست، آموزش و اعتبارسنجی به ترتیب با سهم‌های ۱۵٪، ۷۰٪ و ۱۵٪ تقسیم شدند و بخش دوم داده‌ها (۲۰٪) برای آزمون تکمیلی استفاده شد. برای انتخاب الگوریتم مناسب، آموزش به‌عنوان بخش کلیدی هر فرآیند مدل‌سازی شبکه عصبی مصنوعی، از ۷ الگوریتم آموزشی پس‌انتشار:

Levenberg-Marquardt (trainlm) - conjugate

2. Hyperbolic tangent sigmoid
3. purelin

1. Mean squared error

عصبی و وزن‌های به دست آمده برای آن انجام می‌شود (۱۶).

معادله (۲)

$$I_{ij} = \frac{\sum_{m=1}^{m=Nh} \left(\left(\frac{|W_{jm}^{ih}|}{\sum_{k=1}^{Ni} |W_{km}^{ih}|} \right) \times |W_{mn}^{ho}| \right)}{\sum_{k=1}^{k=Ni} \left\{ \sum_{m=1}^{m=Nh} \left(\frac{|W_{km}^{ih}|}{\sum_{k=1}^{Ni} |W_{km}^{ih}|} \right) \times |W_{mn}^{ho}| \right\}} \times 100$$

در معادله ۲، I_{ij} از اهمیت نسبی (آمین متغیر ورودی بر روی متغیر خروجی، N_k و N_i به ترتیب تعداد نورون‌های لایه ورودی و لایه پنهان هستند، W وزن بین لایه بوده که در آن اندیس‌های i, h, o به ترتیب به لایه‌های ورودی، پنهان و خروجی اشاره دارد. 'n'، 'm' و 'k' به ترتیب تعداد نورون‌ها در لایه‌های خروجی، پنهان و ورودی هستند (۱۶).

ارزیابی مدل‌ها

ضریب همبستگی (R)، میانگین خطای مطلق (MAE) و MSE برای ارزیابی نیکویی برازش و پیش‌بینی دقت مدل‌ها در نظر گرفته شد. به طور کلی، بهینه‌ترین مدل مربوط به مدلی است که مقادیر بالای R و مقادیر کوچک MSE و MAE را دارد. این شاخص‌های آماری در نرم‌افزار متلب به صورت زیر توسعه داده شد (۱۶).

$$R = \text{corr}(\text{Real Turbidity values, anticipated Turbidity values}) \quad (۳)$$

$$MSE = \text{mean}((\text{Real Turbidity values} - \text{anticipated Turbidity values})^2) \quad (۴)$$

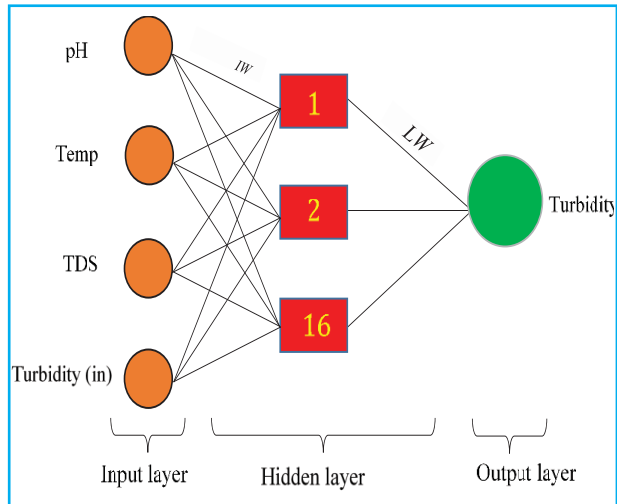
$$MAE = \text{mean}(\text{abs}(\text{Real Turbidity values} - \text{anticipated Turbidity values})) \quad (۵)$$

یافته‌ها

توسعه مدل شبکه عصبی مصنوعی برای کدورت خروجی تصفیه‌خانه آب طبس

انتخاب الگوریتم آموزش پس‌انتشار

برای انتخاب الگوریتم آموزش پس‌انتشار مناسب، ۷ الگوریتم بر اساس جدول ۱ مورد آزمون قرار گرفتند. عملکرد الگوریتم‌های آموزشی پس‌انتشار با در نظر گرفتن حداقل مقدار MSE برای هر یک از نورون‌ها در لایه‌های پنهان شبکه‌های عصبی مصنوعی



شکل ۲. معماری شبکه عصبی مصنوعی توسعه داده شده

رگرسیون خطی چندگانه

در رگرسیون خطی چندگانه، محقق به دنبال پیش‌بینی تغییرات بیش از یک متغیر مستقل بر روی یک متغیر وابسته است. (۱۳). به طور کلی برای برآورد پارامترهای مدل رگرسیون خطی چندگانه، باید از روش ماتریسی استفاده نمود که این روش‌ها را معمولاً با نرم‌افزار انجام داده و پارامترها را برآورد می‌کنند (۱۴). معادله خط رگرسیون چندگانه (معادله ۱) به صورت زیر است:

معادله (۱)

$$Y = a + b_1 X_1 + b_2 X_2 + e_i \quad i = 1, 2, \dots$$

در این معادله a مقدار ثابت، b مقدار ضریب به دست آمده برای هر متغیر و X نوع متغیر است. در این مطالعه از نرم‌افزار متلب برای توسعه مدل رگرسیون خطی چندگانه استفاده شد.

تحلیل حساسیت مدل ANN

تجزیه و تحلیل حساسیت، اهمیت متغیرهای مستقل مانند pH، دما، کدورت ورودی و کل جامدات محلول را در مدل‌های ANN ارزیابی می‌کند. در حقیقت تحلیل حساسیت، روشی است که در آن می‌توان مهم‌ترین متغیرهای ورودی که بیشترین تأثیر بر خروجی مدل دارند را شناسایی کرد (۱۵). برای شناسایی مؤثرترین متغیر مدل ANN، تحلیل حساسیت با تکیه بر معادله گارسون (معادله ۲) اجرا شد. اساس این معادله، با توجه به بهترین مدل شبکه

1. Mean absolute error

توسعه داده شده با این الگوریتم‌ها قابل بررسی و مقایسه است. در این مطالعه با انجام تست‌های اولیه در انتخاب الگوریتم آموزش پس‌انتشار، الگوریتم لونیبرگ-مارکوآرت (trainlm) از میان ۷ الگوریتم مورد بررسی، به‌عنوان بهترین الگوریتم آموزشی شناخته شد. بر این اساس، از الگوریتم لونیبرگ-مارکوآرت در توسعه مدل شبکه عصبی مصنوعی و ارزیابی‌های بعدی استفاده گردید.

جدول ۱. انتخاب الگوریتم آموزشی پس‌انتشار

کدورت		الگوریتم‌های پس‌انتشار		
بهترین نورون لایه پنهان	تعداد تکرار	MSE	R	
۱۲	۳۱	۰/۰۰۶۸۴	۰/۷۵۱۲۱	traincgp (Polak-Ribière conjugate gradient)
۱۷	۱۰۰۰	۰/۰۰۹۵	۰/۷۱۲۳۴	traingdm (gradient descent with momentum)
۲۰	۴۵	۰/۰۰۱۷۱	۰/۸۴۱۲۸	trainscg (scaled conjugate gradient)
۱۴	۳۰	۰/۰۰۱۳۲۱	۰/۸۶۲۱۴	trainrp (resilient backpropagation)
۹	۱۵	۰/۰۰۱۲۱۲	۰/۸۷۱۹۵	trainoss (one step secant)
۱۸	۲۳	۰/۰۰۱۴۵	۰/۸۵۱۶۵	trainbfg (BFGS quasi-Newton)
۱۶	۹	۰/۰۰۰۹۷	۰/۸۹۲۱۳	trainlm (Levenberg-Marquardt)

جدول ۲. بررسی عملکرد نورون‌ها در پیش‌بینی کدورت خروجی تصفیه‌خانه

نورون	MSE مجموع	MSE داده‌های آموزش	MSE داده‌های واقعی	MSE داده‌های آزمایشی
۱	۰/۰۰۶۹۰۳	۰/۰۰۲۴۶	۰/۰۰۲۳۳۸۳۵۶	۰/۰۰۲۱۰۴
۲	۰/۰۰۶۷۰۸	۰/۰۰۲۳۷۱	۰/۰۰۱۷۴۴۰۴۳	۰/۰۰۲۵۹۳
۳	۰/۰۰۷۰۴۶	۰/۰۰۲۲۳۷	۰/۰۰۲۲۵۱۰۷۸	۰/۰۰۲۵۵۸
۴	۰/۰۰۵۴۴۹	۰/۰۰۱۲۶۹	۰/۰۰۲۲۱۷۳۸۹	۰/۰۰۱۹۶۳
۵	۰/۰۰۴۲	۰/۰۰۱۲۲۵	۰/۰۰۲۰۲۸۳۹۲	۰/۰۰۰۹۴۶
۶	۰/۰۰۵۹۹۸	۰/۰۰۲۲۴۵	۰/۰۰۱۴۶۵۵۴۷	۰/۰۰۲۲۸۷
۷	۰/۰۰۵۶۱۳	۰/۰۰۱۲۷۵	۰/۰۰۰۸۰۴۸۹۴	۰/۰۰۳۵۳۲
۸	۰/۰۰۵۹۷۷	۰/۰۰۱۰۵۳	۰/۰۰۳۰۰۹۳۵۳	۰/۰۰۱۹۱۴
۹	۰/۰۰۴۶۹	۰/۰۰۰۹۸۴	۰/۰۰۲۶۱۸۰۴۲	۰/۰۰۱۰۸۸
۱۰	۰/۰۰۴۵۵۸	۰/۰۰۰۷۴۸	۰/۰۰۱۸۰۸۳	۰/۰۰۲۰۰۲
۱۱	۰/۰۰۴۴۸۳	۰/۰۰۰۳۲۵	۰/۰۰۰۸۱۷۵۱۹	۰/۰۰۳۳۴۱
۱۲	۰/۰۰۷۴۷۳	۰/۰۰۱۵۵۴	۰/۰۰۳۵۹۶۱۹۱	۰/۰۰۲۳۲۳
۱۳	۰/۰۰۳۶۲	۰/۰۰۱۳۰۶	۰/۰۰۱۵۱۴۳۰۳	۰/۰۰۰۸
۱۴	۰/۰۰۳۹۲۷	۰/۰۰۰۸۷۳	۰/۰۰۱۵۳۵۱۷۲	۰/۰۰۱۵۱۹
۱۵	۰/۰۰۷۴۸۱	۰/۰۰۰۸۳۷	۰/۰۰۲۳۷۷۵۹۶	۰/۰۰۴۲۶۷
۱۶	۰/۰۰۳۴۸۱	۰/۰۰۰۸۲۲	۰/۰۰۱۰۰۶۴۹۷	۰/۰۰۱۶۵۲
۱۷	۰/۰۰۵۸۸۷	۰/۰۰۰۸۰۶	۰/۰۰۱۶۵۴۸۷۶	۰/۰۰۳۴۲۶
۱۸	۰/۰۰۴۵	۰/۰۰۱۶۳۳	۰/۰۰۱۳۳۱۴۵	۰/۰۰۱۴۸۳
۱۹	۰/۰۰۷۰۴۷	۰/۰۰۲۱۷۹	۰/۰۰۲۲۸۴۴۵۱	۰/۰۰۲۵۸۳
۲۰	۰/۰۰۳۴۹۹	۰/۰۰۰۸۱	۰/۰۰۱۰۷۵۸۶۲	۰/۰۰۱۶۱۳

بهینه‌سازی تعداد نورون مدل شبکه عصبی مصنوعی با الگوریتم لونیبرگ-مارکوآرت

در مدل شبکه عصبی مصنوعی با الگوریتم لونیبرگ-مارکوآرت برای بهینه‌سازی تعداد نورون‌ها، تعداد نورون ۱ تا ۲۰ در لایه پنهان مدل شبکه عصبی در نظر گرفته شد. هر یک از این نورون‌ها ۲۰ بار تکرار شدند، این عمل به این معناست که برای هر یک از نورون‌ها، ۲۰ مدل شبکه عصبی ساخته شده است و مدل‌هایی که کمترین خطا را داشتند، برای مقایسه با یکدیگر لحاظ شدند. اساس کار، حداقل مقدار MSE در سه مرحله اعتبارسنجی، آموزش و آزمایش یا مجموع آنها است. همانطور که در جدول ۲ قابل مشاهده است، کوچک‌ترین MSE در الگوریتم trainlm برای نورون شماره ۱۶ بود، بنابراین این نورون به‌عنوان انتخاب مناسب برای پیش‌بینی کدورت خروجی در نظر گرفته شد. از این رو، معماری ۱-۱۶-۴ (شکل ۲) به نظر می‌رسد بهینه‌ترین توپولوژی برای مدل‌های شبکه عصبی مصنوعی برای پیش‌بینی کدورت خروجی باشد که در این توپولوژی اعداد ۱ و ۴ به ترتیب تعداد نورون‌ها در لایه‌های خروجی و ورودی هستند. نوع توابع انتقال در لایه‌های مخفی و خروجی به ترتیب تابع انتقال tansig و تابع انتقال خطی purelin بودند. محدوده نرمال‌سازی داده‌ها نیز در بازه ۰/۱ تا ۰/۹ انجام گردید (معادله ۶).

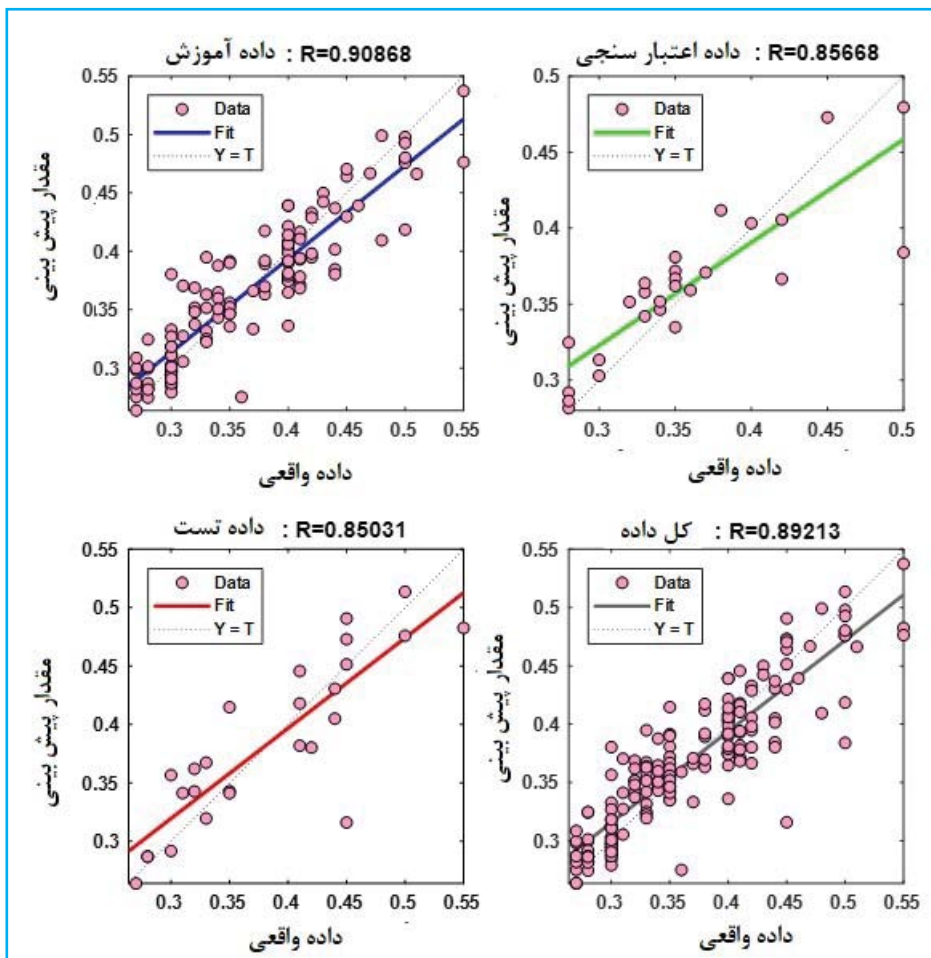
مدل بهینه توسعه داده شده در برابر مقادیر واقعی آن است. ضریب همبستگی به عنوان شاخصی برای همبستگی داده‌های مقادیر پیش‌بینی شده و داده‌های واقعی در نمودار ۱ ذکر شده است. همانطور که در این نمودار مشهود است، مقادیر ضریب همبستگی برای مراحل اعتبارسنجی، آموزش و تست به ترتیب ۰/۸۵۶۶، ۰/۹۰۸۶ و ۰/۸۵۰۳ به دست آمد، در حالی که برای تمام داده‌ها، مقدار R برابر ۰/۸۹۲۱ حاصل شد، بنابراین می‌توان نتیجه گرفت که مدل توسعه یافته می‌تواند مقادیر کدورت خروجی تصفیه‌خانه را به دقت مناسبی پیش‌بینی نماید.

$$\text{معادله (۶)} \quad X_{\text{norm}} = 0.8 \times \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} + 0.1$$

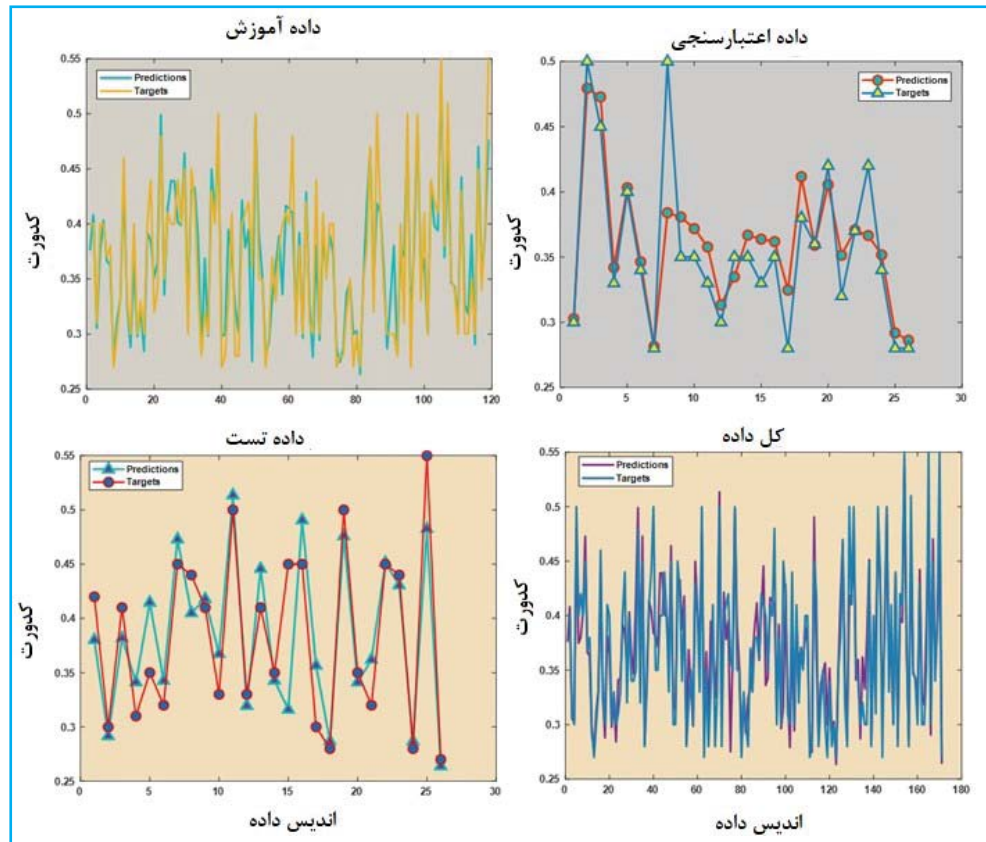
در این معادله، X نشان‌دهنده مقدار اولیه داده‌ها، X_{min} نشان‌دهنده حداقل مقدار داده‌ها و X_{max} نشان‌دهنده حداکثر مقدار داده‌ها است.

اعتبارسنجی و تست مدل

به طور کلی ۳۰٪ از داده‌ها برای مراحل تست و اعتبارسنجی (۱۵٪ سهم هر کدام) مورد استفاده قرار گرفتند. عملکرد پیش‌بینی داده‌ها و مدل ANN با استفاده از MSE و MAE مورد ارزیابی قرار گرفت. نمودارهای ۱ و ۲ نشان‌دهنده مقادیر پیش‌بینی شده کدورت توسط



نمودار ۱. نمودارهای پراکندگی مقادیر پیش‌بینی شده کدورت در برابر مقادیر واقعی نرمال شده



نمودار ۲. مقادیر نرمال شده واقعی و پیش‌بینی کدورت خروجی از تصفیه‌خانه

آمده مابین لایه ورودی و لایه پنهان شبکه عصبی مصنوعی و b_1 و b_2 به ترتیب بایاس‌های اعمال شده به لایه پنهان و لایه خروجی شبکه عصبی مصنوعی است.

همچنین معادله ۸ ارتباط کلی داده‌های واقعی و پیش‌بینی شده کدورت خروجی از تصفیه‌خانه برای داده‌های نرمال شده را نشان می‌دهد.

معادله (۸)

$$\text{Turbidity}_{pred} = 0.78 \times \text{Turbidity}_{Actual} + 0.081$$

برای اینکه قدرت اعتبارسنجی مدل‌های توسعه داده شده بیشتر مورد بررسی قرار گیرد، یک تست اضافی نیز با ۴۴ داده مجزا انجام شد. نمودار ۴ نتایج این مرحله را نشان می‌دهد. بر اساس نمودار ۴، داده‌های پیش‌بینی شده به شکل مناسبی از روند داده‌های واقعی پیروی می‌کنند. از طرفی نمودار رگرسیون توسعه داده شده در این مرحله نشان از ارتباط مناسب بین کدورت واقعی و پیش‌بینی با

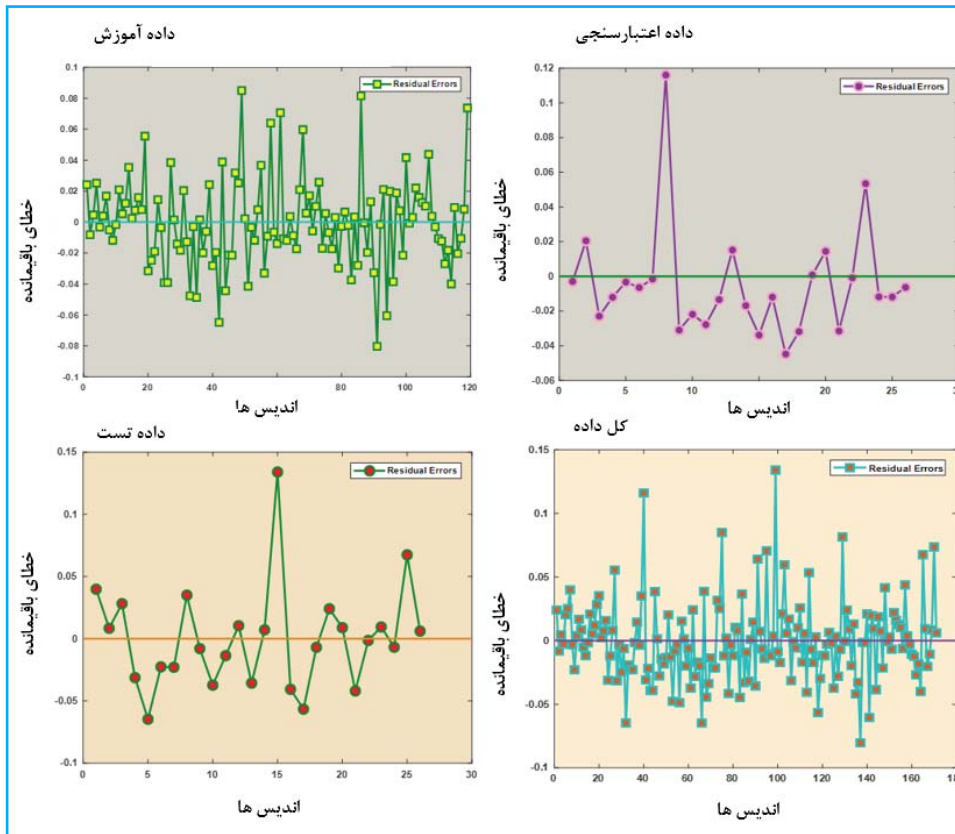
همچنین نتایج مرتبط با خطاهای محاسبه شده در نمودار ۳ نشان داده است. قابل ذکر است که خطای MSE در سه مرحله تست، اعتبارسنجی و آموزش به ترتیب 0.0010 ، 0.0016 ، و 0.0082 به دست آمد (نمودار ۳). علاوه بر این، نسبت میانگین قدر مطلق خطا برای مراحل ذکر شده به ترتیب 0.0217 ، 0.0295 و 0.0213 به دست آمد. مقادیر خطای باقی‌مانده برای هر سه مرحله تست، اعتبارسنجی و آموزش در نمودار ۳ نشان داده شده است.

مدل شماتیک ANN توسعه داده شده برای پیش‌بینی کدورت در مطالعه حاضر در معادله ۷ ارائه شده است:

معادله (۷)

$$\text{ANN equation} = \text{Purelin} \{ W_2 * \text{tansig} (W_1 * [pH; TDS; Temp; [\text{Turbidity}]_{in}] + b_1) + b_2 \}$$

در معادله فوق W_2 وزن ارتباطی به دست آمده بین لایه پنهان و لایه خروجی شبکه عصبی مصنوعی، W_1 وزن ارتباطی به دست



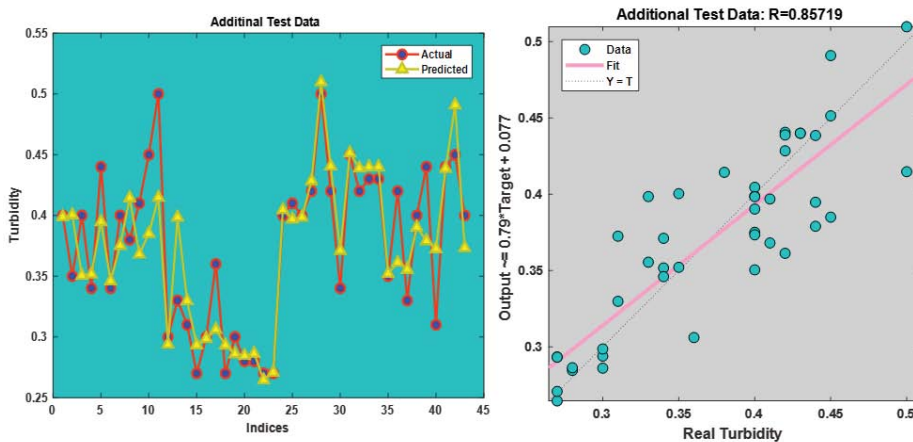
نمودار ۳. مقادیر خطای باقی‌مانده، MSE و MAE در پیش‌بینی کدورت خروجی از تصفیه‌خانه با مدل شبکه عصبی مصنوعی

خطی چندگانه است. بر اساس تحلیل رگرسیون، یک رابطه خطی بین متغیر پاسخ با یک یا چند متغیر توصیفی برقرار می‌شود. البته گاهی به متغیر پاسخ، متغیر وابسته و به متغیرهای توصیفی، متغیرهای مستقل نیز می‌گویند. در رگرسیون خطی چندگانه،

ضریب همبستگی برابر با 0.857 داشت. همچنین مقدار خطای میانگین مربعات برابر با 0.0011 حاصل شد.

رگرسیون خطی چندگانه

یکی از روش‌های مرسوم در تحلیل چندمتغیره، تکنیک رگرسیون



نمودار ۴. نتایج مرحله تست اضافی مدل شبکه عصبی مصنوعی توسعه داده شده

پارامترهای یک مدل خطی به کمک یک تابع هدف و مقدارهای متغیرها، برآورد می‌شوند. به‌طور کلی رگرسیون خطی، یک رابطه خطی برحسب پارامترهای مدل است (۱۳). در این مطالعه برای انجام رگرسیون خطی چندگانه، داده‌ها به دو بخش ۸۰٪ برای آموزش و ۲۰٪ برای تست مورد استفاده قرار گرفت. نتایج ضریب همبستگی برای داده‌های آموزش ۰/۶۳ و برای داده‌های تست ۰/۶۰ به‌دست آمد. همچنین مقدار خطای میانگین مربعات به ترتیب برابر با ۰/۰۰۲۹ و ۰/۰۰۲۷ برای داده‌های آموزش و تست حاصل شد.

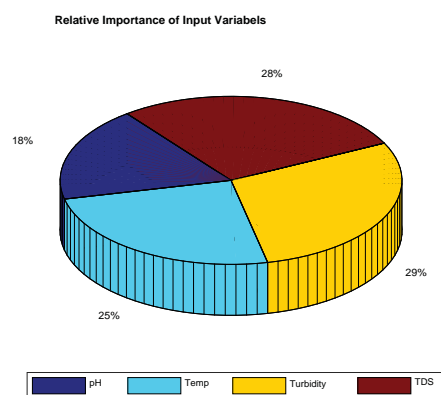
تجزیه و تحلیل میزان حساسیت مدل شبکه عصبی مصنوعی همانطور که در بخش مواد و روش‌ها گفته شد، تجزیه و تحلیل حساسیت، اهمیت متغیرهای مستقل مانند pH، دما، کدورت ورودی و کل جامدات محلول را در مدل‌های ANN ارزیابی می‌کند. نتایج به‌دست آمده برای وزن‌ها و تجزیه و تحلیل حساسیت به ترتیب در جدول ۳ و نمودار ۵ نشان داده شده است.

جدول ۳. وزن اتصال بین لایه‌های مدل شبکه عصبی مصنوعی بهینه

شماره نورون	IW		LW	
	دما	کدورت ورودی	کل جامدات محلول	کدورت خروجی
۱	-۰/۳۱۷۴۹	۱/۶۹۲۵۷۰۰۴۴	۰/۳۵۱۳۶۸	۰/۶۹۴۴۰۰۰۳۴
۲	-۰/۶۰۱۳۵	-۰/۸۴۸۸۴۰۱۴۵	-۲/۱۲۳۵۲	-۰/۹۴۹۷۲۱۸۲
۳	۰/۱۱۱۳۱۱	۲/۱۱۱۹۹۶۰۹۲	-۱/۱۱۵۲۵	-۱/۳۴۸۰۷۳۴۷
۴	۱/۱۴۹۶۷۷	-۱/۰۶۴۱۹۸۷۳۹	۱/۲۲۸۱۸۱	-۱/۶۸۵۴۶۷۴۵۹
۵	-۰/۵۴۵۴۱	-۱/۷۶۰۴۷۴۰۳۶	-۰/۸۲۸۱۴	۰/۴۹۴۰۳۳۳۷۱
۶	-۱/۶۲۳۳۱	۱/۲۲۴۴۲۹۳۶	-۰/۵۶۵۵۶	۰/۲۰۷۸۸۶۵۹۸
۷	-۱/۳۹۷۲	۱/۱۳۳۱۲۸۳۹۶	۰/۸۰۴۵۸۵	۰/۴۹۰۱۸۰۰۶۵
۸	-۰/۰۸۲۴۹	-۰/۴۷۰۲۵۶۶۱۵	۰/۵۷۶۷۲۷	-۱/۷۳۲۰۵۲۴۴۶
۹	-۱/۶۰۵۷۳	۰/۳۴۳۵۲۹۷۱۶	-۱/۶۹۷۸۴	-۰/۳۵۲۰۸۹۰۰۶
۱۰	-۱/۰۵۹۸۱	-۰/۵۶۱۷۲۲۹۱۸	۱/۱۳۲۷۱۵	۱/۳۲۵۶۸۹۸۴
۱۱	۰/۱۸۸۷۰۲	۲/۳۱۲۹۴۸	-۰/۱۵۷۵۸	-۰/۶۵۷۵۳۴۴۱۲
۱۲	-۱/۶۰۶۱۶	-۱/۷۷۶۳۰۷۵۳۶	۰/۳۳۹۶۸۶	-۱/۰۳۸۸۱۱۸۱۱
۱۳	-۰/۳۰۲۷۵	۱/۱۵۰۴۹۸	-۱/۶۳۰۰۲	۱/۳۹۴۳۸۰۶۸۶
۱۴	-۰/۳۳۵۹	-۰/۴۴۴۷۶	۱/۱۶۰۶۲۱۰۷۶	-۱/۱۵۸۹۲۹۶۴۴
۱۵	۰/۲۸۱۳۹۵	۱/۱۶۷۳۵۷	۱/۲۱۶۵۵۲	۰/۰۲۴۰۸۵۷۵۳
۱۶	۰/۲۶۴۲۹۱	-۱/۱۰۳۵۱	۱/۵۸۹۱۲	-۰/۷۵۷۹۴۵۳۱

بحث

با توجه به نتایج به‌دست آمده در مطالعه حاضر (مقادیر بالاتر ضریب رگرسیون و متعاقباً مقادیر پایین خطا)، مدل شبکه عصبی مصنوعی بر مدل رگرسیون خطی چندگانه برتری داشت. خدیر و همکاران مطالعه‌ای با عنوان ارزیابی عملکرد تصفیه‌خانه آب AI-Karkh با استفاده از مدل‌های داده محور و مدل محور انجام دادند که این مطالعه متمرکز بر مدل‌سازی ریاضی فرآیندهای تصفیه آب با انگیزه اولیه برای تهیه ابزاری بود که می‌تواند برای پیش‌بینی



نمودار ۵. نتایج آنالیز حساسیت مدل شبکه عصبی مصنوعی توسعه داده شده

عملکرد تصفیه جهت کنترل بهتر عدم قطعیت استفاده شود. طبق این مطالعه، مشخص شد که از پارامترهای مهم آب خام: سختی کل، کلسیم، منیزیم، جامدات کل، نیتريت، نترات، آمونیاک و سیلیس برای ساخت مدل استفاده شود. برای مدل‌سازی دمای فرآیند انعقاد و لخته‌سازی، قلیابیت و pH آب خام، متغیرهای وابسته مدل بودند. در مورد مدل‌سازی کدورت آب تصفیه شده، متغیرهای ورودی مدل شامل: غلظت‌های آهن، قلیابیت، هدایت الکتریکی، دما، منیزیم، کلسیم، کل جامدات، سولفات، سختی کل و کلراید بودند و کدورت آب تصفیه شده به‌عنوان متغیر خروجی استفاده شده بود. بر اساس نتایج این مطالعه، مدل‌های رگرسیون خطی چندمتغیره و رگرسیون خطی چندگانه دارای عملکرد پیش‌بینی کمی بالاتر از مدل داده محور شبکه عصبی مصنوعی بودند. مقادیر ضریب همبستگی در این مطالعه برای مدل‌های رگرسیون خطی چندگانه و شبکه عصبی مصنوعی به ترتیب $0/831$ و $0/772$ بود (۱۷). سلیم و الموسوی به بررسی عملکرد تصفیه‌خانه آب Al-Wahda با استفاده از مدل‌های رگرسیون خطی چندگانه و شبکه عصبی مصنوعی پرداختند. مقادیر ضریب همبستگی در این مطالعه برای مدل‌های رگرسیون خطی چندگانه و شبکه عصبی مصنوعی به ترتیب $0/887$ و $0/925$ بود. این محققان نتیجه گرفتند که مدل‌های شبکه عصبی مصنوعی بر رگرسیون خطی چندگانه در پیش‌بینی کدورت خروجی از تصفیه‌خانه برتری دارد (۱۰). با توجه به نتایج به‌دست آمده در این تحقیق، مطالعه حاضر با مطالعه سلیم و الموسوی تطابق بیشتری داشت. البته این نکته نیز قابل ذکر است که خدیر و همکاران احتمالاً بررسی جامع‌تری بر روی مدل شبکه عصبی مصنوعی انجام ندادند و یا اینکه تجربیات و مهارت لازم را در توسعه مدل شبکه عصبی نداشتند؛ چراکه مدل‌های شبکه عصبی مصنوعی قادر به دریافت خصوصیات و روابط خطی و غیرخطی بین متغیرها هستند. در مطالعه حاضر تجزیه و تحلیل حساسیت در مدل شبکه عصبی مصنوعی توسعه داده شده برای بررسی اهمیت متغیرهای مستقل مختلف (pH و دما، کدورت ورودی و کل جامدات محلول)

در خروجی مدل انجام گرفت. این رویکرد توسط وزن اتصال بین لایه‌های مدل شبکه عصبی مصنوعی انجام می‌گیرد. جدول ۳ ماتریس وزن شبکه عصبی را برای مدل‌سازی کدورت خروجی تصفیه‌خانه آب طبس نشان می‌دهد. بر اساس نتایج به‌دست آمده از تحلیل گارسون در انجام آنالیز حساسیت مشخص می‌شود که تمام متغیرها نقش اساسی در پیش‌بینی کدورت خروجی تصفیه‌خانه داشتند. به‌طور کلی، حداکثر اثربخشی متعلق به کدورت ورودی با مقدار 29% است. در حالی که حداقل تأثیر مربوط به pH با درصد اهمیت برابر با 18% بود.

نتیجه‌گیری

بر اساس نتایج مطالعه حاضر، مدل شبکه عصبی مصنوعی در مقایسه با رگرسیون خطی چندگانه در پیش‌بینی کدورت خروجی تصفیه‌خانه آب طبس نتایج بهتری را فراهم نموده؛ در نتیجه به‌عنوان بهترین مدل انتخاب شد. بر اساس نتایج تحلیل حساسیت بر طبق معادله گارسون، مشخص شد که کدورت ورودی با مقدار 29% به‌عنوان مهم‌ترین پارامتر در توسعه مدل شبکه عصبی مصنوعی است. علاوه بر این بر اساس شاخص‌های آماری مشخص شد که شبکه عصبی مصنوعی بر اساس الگوریتم آموزشی لونیبرگ-مارکوآرت، قابلیت پیش‌بینی با دقت بالاتری را نسبت به سایر مدل‌های شبکه عصبی مصنوعی پس‌انتشار برای پیش‌بینی مقادیر کدورت خروجی تصفیه‌خانه دارد.

ملاحظات اخلاقی

نویسندگان تمام نکات اخلاقی شامل سرقت ادبی، انتشار دوگانه، تحریف داده‌ها و داده‌سازی را در این مقاله رعایت کرده‌اند. همچنین هرگونه تضاد منافع حقیقی یا مادی که ممکن است بر نتایج یا تفسیر مقاله تأثیر بگذارد را رد می‌کنند.

تشکر و قدردانی

مطالعه حاضر تحت عنوان پروژه تحقیقاتی دانشجویی به شماره طرح IR.BUMS.REC.1400.255 معاونت پژوهشی دانشگاه علوم پزشکی بیرجند انجام شده است. بدین وسیله از تمامی اساتیدی که ما را در نگارش این پژوهش یاری نمودند، تقدیر و تشکر می‌گردد.

References

1. Twort AC, Ratnayaka DD, Brandt MJ. Water supply: Elsevier; 2000.
2. Organization WH. Water quality and health-review of turbidity: information for regulators and water suppliers. 2017.
3. Inspectorate DW. Guidance on the implementation of the water supply (water quality) regulations 2000 (as amended) in England. Drinking Water Inspectorate Guidance Document. 2012.
4. Stevenson M, Bravo C. Advanced turbidity prediction for operational water supply planning. *Decision Support Systems*. 2019;119:72-84.
5. Goerger A, Darmaillacq AS, Shashar N, Dickel L. Early Exposure to Water Turbidity Affects Visual Capacities in Cuttlefish (*Sepia officinalis*). *Frontiers in physiology*. 2021;12:622126.
6. Cleary SA. Sustainable Drinking Water Treatment for Small Communities Using Multistage Slow Sand Filtration: UWSpace; 2005.
7. matuzahroh N, Fitriani N, Ardiyanti PE, Kuncoro EP, Budiyanto WD, Isnadina DRM, et al. Behavior of schmutzdecke with varied filtration rates of slow sand filter to remove total coliforms. *Heliyon*. 2020;6(4):e03736.
8. Ranjan P, Prem M. Schmutzdecke-a filtration layer of slow sand filter. *International Journal of Current Microbiology and Applied Sciences*. 2018;7(07):637-45.
9. Alade O, Al Shehri D, Mahmoud M, Sasaki K. Viscosity–Temperature–Pressure Relationship of Extra-Heavy Oil (Bitumen): Empirical Modelling versus Artificial Neural Network (ANN). 2019;12(12):2390.
10. Saleem AT, Al-Musawi NO. Evaluation the turbidity removal efficiency in Al-Wahda water treatment plant using statistical indicators. *Journal of Physics: Conference Series*. 2021;1895(1):012045.
11. Kriegeskorte N, Golan T. Neural network models and deep learning. *Current Biology*. 2019;29(7):R231-R6.
12. Al-Musawi N. Application of Artificial Neural Network for Predicting Iron Concentration in the Location of Al-Wahda Water Treatment Plant in Baghdad City. *University of Baghdad Engineering Journal*. 2018;22:72-82.
13. Hosseinzadeh A, Baziar M, Alidadi H, Zhou JL, Altaee A, Najafpoor AA, et al. Application of artificial neural network and multiple linear regression in modeling nutrient recovery in vermicompost under different conditions. *Bioresource Technology*. 2020;303:122926.
14. Marill KA. *Advanced Statistics: Linear Regression, Part II: Multiple Linear Regression*. 2004;11(1):94-102.
15. Shojaeefard MH, Akbari M, Tahani M, Farhani F. Sensitivity Analysis of the Artificial Neural Network Outputs in Friction Stir Lap Joining of Aluminum to Brass. *Advances in Materials Science and Engineering*. 2013;2013:574914.
16. Mahmood Yousefi AE, Mitra Gholami, Ali Akbar Mohammadi, Nadeem A. Khan, Mansour Baziar, Vahide Oskoei,. Modeling and analysis of the groundwater hardness variations process using machine learning procedure. *Desalination and Water Treatment*. 2021;238:7.
17. Khudhair BH, Alsaqqar AS, Jbbar RK. Performance Evaluation of Al-Karkh Water Treatment Plant Using Model-driven and Data-Driven Models. *IOP Conference Series: Earth and Environmental Science*. 2021;779(1):012110.