

به کارگیری روش های آماری برای افزایش دقت مدل سازی تصفیه خانه های فاضلاب صنعتی با استفاده از شبکه های عصبی مصنوعی

میرحسین سید سراجی^{۱*}، حامد حسنلو^۲، مریم پازوکی^۲، حسین نایب^۳

۱. تهران، دانشگاه شهید بهشتی، دانشکده آب و محیط زیست، گروه مهندسی منابع آب

۲. تهران، دانشگاه تهران، دانشکده محیط زیست، گروه مهندسی محیط زیست

۳. تهران، دانشگاه شهید بهشتی، دانشکده آب و محیط زیست، گروه مهندسی محیط زیست

(تاریخ دریافت: ۹۳/۰۴/۱۱ - تاریخ پذیرش: ۹۴/۰۳/۰۱)

چکیده

به منظور کنترل بهتر و کارآمد عملکرد تصفیه خانه های فاضلاب صنعتی، می توان از ابزاری ریاضی برمبنای اطلاعات ثبت شده برخی از پارامترهای اساسی پساب، طی دوره ای از بهره برداری تصفیه خانه استفاده کرد. در این پژوهش، برای اولین بار در کشور از شبکه عصبی چندلایه پیش خور با یک لایه پنهان و روش توقف آموزش، به منظور بررسی مشخصات پساب خروجی واحدهای تصفیه خانه استفاده شده است. همچنین، از روش تحلیل عاملی برای اصلاح و ارتقای عملکرد مدل های ترکیبی ایجاد شده از طریق شبکه عصبی و تکنیک تحلیل مؤلفه های اصلی استفاده شده است. این تحلیل یکی از روش های آماری است که برای تجزیه اطلاعات موجود در مجموعه داده ها و تعیین تأثیرگذارترین متغیرها در هنگام زیاد بودن تعداد متغیرهای مورد بررسی و ناشناخته بودن روابط بین آنها استفاده می شود. به منظور ارزیابی عملکرد مدل ها از شاخص های مجذور میانگین مربع خطاها (RMSE)، میانگین مطلق خطا (MAE) و ضریب همبستگی پیرسون (R) استفاده شده است. مقادیر R به دست آمده از مدل ها که در بازه ۰/۸ تا ۰/۹۴ قرار دارند، نشان دهنده دقت مناسب آن در برآورد مشخصات کیفی فاضلاب است.

واژه های کلیدی: پتروشیمی فجر، تحلیل عاملی، تحلیل مؤلفه اصلی، تصفیه خانه فاضلاب صنعتی، شبکه عصبی مصنوعی.

۱. مقدمه و هدف

مطالعه موردی این پژوهش انتخاب شده است، از جمله تصفیه خانه های ویژه به شمار می رود، زیرا پساب ورودی به تصفیه خانه در درجه اول حاوی نمک های محلول (تی دی اس)^۱ بسیار بالاست و در ضمن گستره تغییرات این مشخصه از پساب نیز زیاد است و کوچک ترین اشتباه در امر راهبری تصفیه خانه و نبود تصفیه مناسب تا حد استاندارد تعیین شده، موجب مشکلات جدی زیست محیطی در منطقه ویژه می شود. همچنین، خارج شدن تعدادی از واحدهای تصفیه به دلیل عوامل مختلف از جمله سرویس های دوره ای، وقوع حادثه و ورود پساب غیرمعمول

فرایندهای موجود در سیستم های محیط زیستی که مهندسان محیط زیست با آن مواجه اند، اغلب دو خصوصیت عمده دارند: ۱. به متغیرهای زیادی وابسته اند؛ ۲. روابط پیچیده ای بین اجزای آن وجود دارد که تحلیل آن را بسیار مشکل می کند [۱]. بهره برداری از تصفیه خانه های فاضلاب صنعتی اغلب به دلیل تنوع بسیار، تغییرات گسترده در کیفیت پساب ورودی به تصفیه خانه و ماهیت طبیعی فرایند تصفیه بیولوژیکی مورد استفاده، پیچیده و دشوار است. تصفیه خانه فاضلاب صنعتی فجر که به عنوان

«مدل سازی شبکه عصبی مصنوعی برای حذف کادمیوم از محیط های آبی» استفاده کرد. در این پژوهش، با ترکیب شبکه عصبی پیش خور و تکنیک تحلیل مؤلفه های اصلی، میزان بازده جذب کادمیوم از سوی دانه های جاذب پیش بینی شد و از مدل عصبی با دو لایه پنهان و تابع آموزش لونیگ-مارکوویت استفاده شد. نتایج این مطالعه نشان دهنده عملکرد مناسب روش ترکیبی شبکه عصبی مصنوعی و تکنیک تحلیل مؤلفه های اصلی است [۵]. در زمینه تحقیقات صورت گرفته درباره شبیه سازی فرایندهای تصفیه فاضلاب، می توان به پژوهش تزل در سال ۲۰۱۰ با عنوان «تهیه مدل عصبی مصنوعی برای کنترل تصفیه خانه فاضلاب شهری» اشاره کرد، که برخلاف پژوهش های قبلی در این زمینه (که عموماً تصفیه خانه ها به عنوان یک بخش واحد در نظر گرفته می شوند)، تصفیه خانه به دو بخش تقسیم می شود و شامل تصفیه اولیه و فرایند بیولوژیک (تصفیه ثانویه) است [۶].

در زمینه کاربرد شبکه های عصبی مصنوعی در تصفیه فاضلاب، می توان به مدل سازی فرایند اکسیداسیون پیشرفته [۷]، مدل سازی آماری راندمان حذف اکسیژن مورد نیاز واکنش شیمیایی (سی او دی) [۴] از پساب صنعت نساجی با استفاده از راکتور یو آ اس بی [۸] و پیش بینی پارامترهای پساب تصفیه خانه های فاضلاب [۹] اشاره کرد. همچنین، حسنلو و همکاران در سال ۲۰۱۲ از روش ترکیبی شبکه عصبی و تکنیک تحلیل مؤلفه های اصلی برای شبیه سازی تصفیه خانه پساب صنعتی استفاده کردند. در تحقیق آن ها، پارامترهای تی دی اس، سی او دی، پی اچ^۷، یون کلر و کدورت مدل شدند [۱۰، ۱۱].

تحلیل مؤلفه های اصلی، تکنیکی است که برای تبدیل متغیرهای متعامد و کاهش ابعاد آن ها به منظور اصلاح و ارتقای عملکرد مدل های ایجاد شده در شبکه عصبی استفاده می شود. کاربرد این روش راه را برای شبکه عصبی در زمینه ایجاد مدلی به منظور شبیه سازی فرایند تصفیه بیولوژیکی هموارتر می کند [۱۲]. تحقیقات صورت گرفته نشان می دهد کاربرد همزمان تکنیک تحلیل مؤلفه های اصلی و شبکه عصبی مصنوعی، نسبت به استفاده مجزای آن، نتایج بهتر و دقیق تری دارد [۴].

طبق توضیحات ذکر شده و براساس اینکه در کارهای تحقیقاتی صورت گرفته فقط از تعداد محدودی از پارامترهای

بر اثر رعایت نکردن استاندارد تخلیه از سوی واحدهای مستقر در منطقه، شرایطی را فراهم می کند که مشخصات پساب در واحدهای مختلف تصفیه خانه با حالت عادی آن ها تفاوت دارد. حال اگر بتوان به طریقی از مشخصات پساب در هر واحد از تصفیه خانه آگاهی یافت، می توان اقدامات لازم را به نحو احسن برای رفع مشکل اتخاذ کرد. مدل سازی سنتی فرایندهای بیولوژیکی مورد استفاده، بر مبنای نوشتن معادلات تعادل برای سرعت رشد میکروارگانیسم ها، مصرف سوبسترا و تشکیل محصولات بوده است. در نتیجه، با توجه به اینکه واکنش های میکروبیولوژیکی به صورت غیرخطی و وابسته به زمان هستند و ماهیت پیچیده ای دارند، استفاده از این گونه مدل سازی ها محدودیت هایی دارد [۲]. مدل سازی یکی از ابزارهای قوی ریاضی است که در حالت کلی به دو روش انجام می گیرد [۳]:

- روش های عددی یا قطعی^۱

- روش های آماری^۲

به طور عمومی روش های آماری به دو گروه کلاسیک (رگرسیون چندمتغیره) و پیشرفته تقسیم می شود. تحقیقات صورت گرفته در زمینه پیش بینی مشخصات پساب با استفاده از روش های آماری، سودمندی این روش ها را به اثبات رسانده است [۳]. شبکه های عصبی مصنوعی از روش های پیش بینی آماری پیشرفته اند که امروزه کاربرد گسترده ای در تمام زمینه های علمی دارند. استفاده از شبکه عصبی که بتواند روابط غیرخطی و پیچیده بین ورودی و خروجی ها را پیش بینی کند، راه حل مناسبی برای جایگزینی با روش رگرسیون خطی چندمتغیره است. در ادامه، سعی شده است اطلاعات کاملی در زمینه استفاده از شبکه عصبی در تحقیقات قبلی ارائه شود.

مک کلاچ و پیتز برای اولین بار فن شبکه عصبی مصنوعی را در سال ۱۹۴۳ مطرح کردند که با وجود به کارگیری ساختمان ساده از این مدل، به سرعت و قدرت محاسباتی آن به شدت توجه شد [۴]. اولیویرا در سال ۲۰۰۲ در تحقیقی با عنوان «شبیه سازی تصفیه خانه های فاضلاب با استفاده از شبکه عصبی مصنوعی و تکنیک تحلیل مؤلفه های اصلی»، با استفاده از ترکیب روش های آماری تحلیل مؤلفه های اصلی و شبکه عصبی مصنوعی، میزان اکسیژن مورد نیاز بیوشیمیایی (بی او دی)^۳ در خروجی تصفیه خانه را پیش بینی کرد [۴]. او در این تحقیق پارامتر بی او دی پساب واحد بیولوژیک تصفیه خانه را مطالعه کرد. کاردام در سال ۲۰۱۰ از

1. Deterministic methods

4. Chemical Oxygen Demand (COD)

2. Statistics methods

5. Upflow Anaerobic Sludge Blanket (UASB)

3. Biological Oxygen Demand (BOD)

7. pH

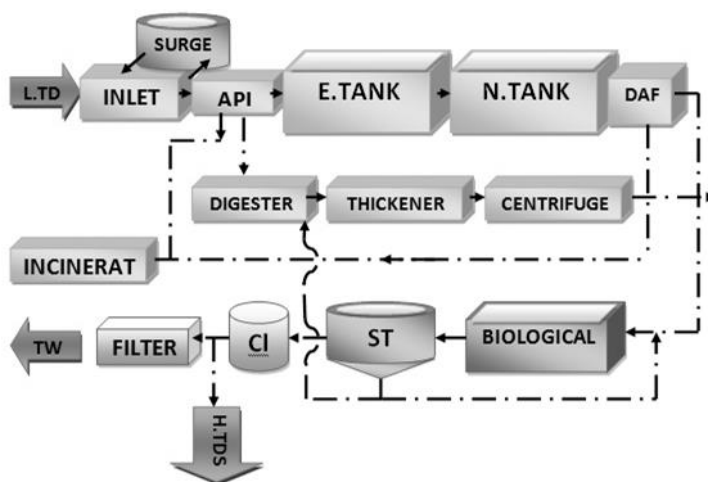
نیاز مجتمع‌های منطقه ویژه اقتصادی پتروشیمی بندر امام خمینی به‌صورت متمرکز احداث شد. در این مجتمع، واحدهای تصفیه پساب ۱ و ۲ به‌منظور تصفیه پساب‌های صنعتی و بهداشتی مجتمع‌های منطقه ویژه احداث شده‌اند [۱۳]. انواع پساب‌های ارسالی از مجتمع‌های منطقه به واحدهای تصفیه پساب پتروشیمی فجر عبارت‌اند از: پساب روغنی با نمک پایین، پساب شیمیایی با نمک بالا، پساب بهداشتی و ضایعات هیدروکربنی. واحد تصفیه پساب در تصفیه‌خانه فجر دارای دو بخش است: ۱. قسمت تصفیه پساب‌های روغنی با نمک پایین و فاضلاب بهداشتی و ۲. قسمت تصفیه پساب شیمیایی با نمک بالا. در این پژوهش، فقط واحد پساب با نمک پایین بررسی شده است. شکل ۱ شماتیک واحد تصفیه با نمک پایین تصفیه‌خانه را نشان می‌دهد. به‌دلیل ثبت و برداشت اطلاعات به‌صورت ساعتی و سه نوبت در روز، داده‌های مورد استفاده در این تحقیق مربوط به بازه بهره‌برداری یک‌ساله تصفیه‌خانه هستند که مقادیر حداکثر، میانگین و حداقل آن‌ها در جدول ۱ ارائه می‌شود. از آنجا که دما به‌جز برداشت‌های واحدها نبود، به‌منظور در نظر گرفتن تأثیر تغییرات دما نیز بازه یک سال (شامل چهار فصل) انتخاب شد.

کیفی به‌عنوان ورودی و خروجی در شبیه‌سازی فرایندهای تصفیه‌خانه استفاده شده است، در این تحقیق به مدل‌سازی واحد تصفیه پساب با نمک پایین تصفیه‌خانه صنعتی برای نخستین بار در کشور توجه شده است و محققان کوشیده‌اند نقص‌های کارهای قبلی تا حد امکان برطرف و راهکارهای جدیدی برای افزایش دقت مدل‌سازی ارائه شود. در نتیجه، هدف این پژوهش ایجاد مدلی است که بتواند در کمترین زمان و با دقت مناسب، مشخصات کیفی پساب خروجی واحدهای تصفیه را ارائه دهد. در این پژوهش، از روش ترکیبی شبکه عصبی و تکنیک تحلیل مؤلفه‌های اصلی برای مدل‌سازی بخش‌های داخلی واحد تصفیه پساب با نمک پایین تصفیه‌خانه فجر استفاده شده است و برای افزایش دقت مدل، از روش تحلیل عاملی به‌منظور تعیین داده‌های ورودی مؤثر به مدل استفاده شده است. پارامترهای استفاده‌شده برای مدل‌سازی شامل سی او دی، پی اچ، تی دی اس، یون کلر و کدورت هستند.

۲. مواد و روش‌ها

۱.۲. منطقه مورد مطالعه (تصفیه‌خانه فاضلاب صنعتی فجر)

شرکت پتروشیمی فجر در سال ۱۳۷۷ با هدف تأمین مواد مورد



L. TDS:	نمک پایین
INLET:	ورودی
SURGE:	مخزن ذخیره (۶۸۰۳)
API:	حوضچه‌های جداکننده روغن (۶۸۱۴)
E. TANK:	مخزن متعادل‌ساز (۶۸۰۲)
N. TANK:	مخزن خنثی‌ساز (۶۸۱۸)
DAF:	جداکننده روغن (۶۸۰۴)
DIGESTER:	هاضم

THICKENER:	تغلیظ‌کننده
CENTRIFUGE:	سانتریفوژ (۶۸۱۰)
INCINERAT:	آبیاری محوطه
TW:	پساب تصفیه‌شده (۶۸۰۵)
FILTER:	فیلتر (۶۸۱۶)
CL:	کلرزنی (۶۸۱۵)
ST:	ته‌نشینی ثانویه (۶۸۱۳)
BIOLOGICAL:	واحد بیولوژیک (۶۸۰۶)

شکل ۱. نمودار فرایند تصفیه پساب تصفیه‌خانه فجر

جدول ۱. مقادیر حداکثر، متوسط و حداقل داده‌های اولیه مورد استفاده در مدل‌سازی

مدل	پارامتر	حداکثر	میانگین	حداقل
L.TDS ورودی واحد	COD (mg/l)	۳۸۵۰	۱۰۱۸/۳۳	۱۸۰
	pH	۱۲	۶/۹	۶/۲
	Oil (mg/l)	۷۵۶	۱۸۰/۶۷	۵۰
	Phenol (mg/l)	۱۳/۴	۱۰/۸۷	۰/۱
	TDS (mg/l)	۲۹۱۰	۵۱۲/۳۳	۴۵۹
L.TDS	pH	۸/۴	۷/۲	۶/۵
	TDS (mg/l)	۲۸۱۰	۱۰۹۸/۱۱۹	۴۱۵
	Turbidity N.T.U	۱۲۱	۵۹/۷۲	۳۰/۶
واحد بیولوژیکی	pH	۷/۹	۷/۲۸	۶/۹
	COD (mg/l)	۴۳۸	۸۷۵/۱۹۸	۵۰
	N (mg/l)	۲۰/۸	۸/۳۷	۲/۲
	P (mg/l)	۶/۴	۱/۶۲	۰/۵

$$KMO = \frac{\sum \sum r_{ij}^2}{\sum \sum r_{ij}^2 + \sum \sum a_{ij}^2} \quad (1)$$

که در آن r_{ij} ضریب همبستگی ساده بین متغیرهای i و j و a ضریب همبستگی جزئی بین آن‌هاست. اگر مجموع ضرایب همبستگی جزئی بین همه زوج متغیرها در مقایسه با مجموع مجذورات ضرایب همبستگی کوچک باشد، اندازه کی ام او نزدیک به یک خواهد بود. مقادیر کوچک کی ام او بیانگر آن است که همبستگی بین زوج متغیرها نمی‌تواند از طریق متغیرهای دیگر تبیین شود؛ بنابراین، کاربرد تحلیل عاملی متغیرها ممکن است قابل توجیه نباشد. جدول ۲ نشان‌دهنده مقادیر قابل قبول ضریب کی ام او است. متغیرهایی که ضریب کی ام او قابل قبول داشته باشند، به‌عنوان داده‌های ورودی مؤثر انتخاب می‌شوند. برای انجام دادن این تحلیل و تعیین متغیرهای مناسب از نرم‌افزار SPSS استفاده شده است. SPSS از جمله نرم‌افزارهای کاربردی در کارهای آماری است. به‌منظور آشنایی در زمینه نحوه کار با این نرم‌افزار و مراحل انجام دادن تحلیل عاملی، به مرجع [۱۴] رجوع شود.

۲.۲ روش تحقیق

۱.۲.۲ روش تحلیل عاملی

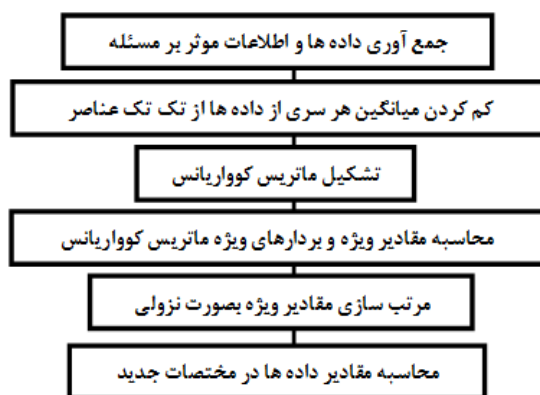
یکی از روش‌های آماری برای تجزیه اطلاعات موجود در مجموعه داده‌ها، روش تجزیه عامل‌ها یا تحلیل عاملی است. این روش برای داده‌هایی رایج و کاربردی است که ابعاد زیادی دارند. پیرسون برای اولین بار روش مذکور را در سال ۱۹۰۱ و سپس اسپیرمن در سال ۱۹۰۴ هنگام اندازه‌گیری هوش مطرح کردند. این روش برای تعیین تأثیرگذارترین متغیرها هنگام زیادبودن تعداد متغیرهای مورد بررسی و ناشناخته‌بودن روابط بین آن‌ها استفاده می‌شود [۱۴]. درکل، هدف تجزیه عامل‌ها شامل موارد زیر است: تفسیر وجود همبستگی درونی بین تعدادی صفت قابل مشاهده از طریق عوامل غیرقابل مشاهده که آن‌ها را عامل گویند (درواقع این عوامل غیرقابل مشاهده دلیل مشترک همبستگی بین متغیرهای اصلی هستند)، ارائه روش ترکیب و خلاصه کردن تعداد زیادی از متغیرها در تعدادی گروه متمایز و تعیین تأثیرگذارترین متغیرها [۱۴]. یکی از روش‌های انتخاب متغیرهای مناسب برای تحلیل عاملی استفاده از ماتریس همبستگی است. از آنجاکه روش تحلیل عاملی بر همبستگی بین متغیرها البته از نوع غیرعالی استوار است، در استفاده از این روش باید ماتریس همبستگی بین متغیرها نیز محاسبه شود. آماره‌های دیگری نیز وجود دارند که پژوهشگر از طریق آن‌ها می‌تواند مناسب بودن داده‌ها را برای تحلیل عاملی تعیین کند. یکی از این روش‌ها استفاده از ضریب کی ام او^۲ است که مقدار آن همواره بین صفر و یک در نوسان است و از رابطه ۱ به‌دست می‌آید:

به‌ویژه زمانی که بیش از دو بعد دارند- دشوار است و نمی‌توان به‌صورت گرافیکی از رابطه بین آن‌ها آگاهی یافت؛ بنابراین، استفاده از این روش برای تحلیل مسائلی مفید است که عوامل زیادی در زمینه مسئله مورد مطالعه دخیل‌اند. تمام داده‌هایی که قرار است به‌عنوان ورودی مدل‌های عصبی استفاده شوند، ابتدا با استفاده از این روش تغییر مختصات داده می‌شوند، سپس وارد مدل‌ها می‌شوند. در ادامه، مراحل مختلف تکنیک تحلیل مؤلفه‌های اصلی ارائه می‌شود [۱۰، ۱۱] (شکل ۲).

مقدار KMO	تناسب داده‌ها با تحلیل عاملی
بزرگ‌تر یا مساوی ۰/۹	عالی
۰/۸۹-۰/۸۰	خیلی خوب
۰/۷۹-۰/۷۰	خوب
۰/۶۹-۰/۶۰	متوسط
۰/۵۹-۰/۵۰	ضعیف
کمتر از ۰/۵	غیرقابل پذیرش

۲.۲.۲. تحلیل مؤلفه‌های اصلی^۱

این روش برای داده‌هایی رایج و کاربردی است که ابعاد زیادی دارند. تشخیص الگویی که داده‌ها از آن پیروی می‌کنند-



شکل ۲. مراحل تکنیک تحلیل مؤلفه‌های اصلی

متغیرها برای استفاده از آن‌ها در شبکه عصبی به روش‌های متفاوتی صورت می‌گیرد. در این پژوهش، از روش صحت‌سنجی موازی استفاده می‌شود. صحت‌سنجی موازی، روشی است که به دفعات در مدل‌های شبکه عصبی استفاده شده است [۱۶]. این روش برای تعیین زمان پایان آموزش و مقایسه توانایی تعمیم مدل‌های متفاوت استفاده می‌شود [۱۷].

۰.۳.۲.۲. شبکه‌های پرسپترون چندلایه

بر اساس اینکه شبکه‌های عصبی ساخته شده با یک نورون، توانایی پیاده‌سازی توابع غیرخطی را ندارند، از شبکه‌های پرسپترون چندلایه جلودرنده با روال تعلیم انتشار به عقب در شبیه‌سازی فرایند بیولوژیکی استفاده شده است. در این نوع شبکه، ورودی‌های لایه اول نرون‌ها به لایه‌های بعدی متصل است و در هر سطح این مسئله صدق می‌کند تا به لایه خروجی برسد. روال انتشار به عقب به معنی این است که پس از مشخص شدن خروجی شبکه، ابتدا وزن‌های لایه آخر تصحیح می‌شوند و به ترتیب اوزان لایه‌های قبلی تصحیح

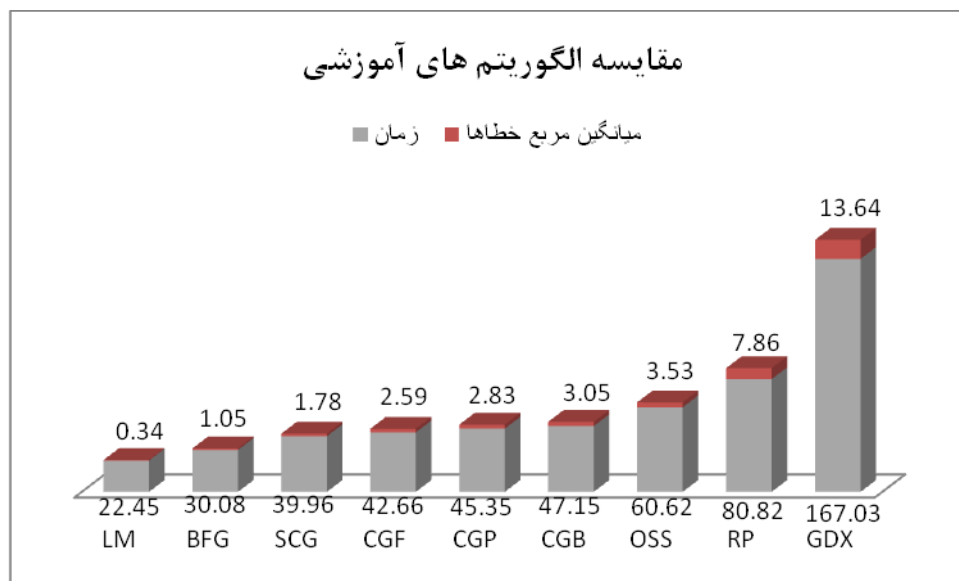
۳.۲.۲. شبکه عصبی مصنوعی^۲

ای‌ان‌ان‌ها مدل‌هایی محاسباتی هستند که می‌توانند رابطه بین ورودی‌ها و خروجی‌های یک دستگاه فیزیکی (هرچند پیچیده و غیرخطی) را با شبکه‌ای از گره‌های متصل به هم تعیین کنند. از جمله مهم‌ترین عوامل برای تعریف ای‌ان‌ان، ساختار معماری آن است. ساختار ای‌ان‌ان به شکلی است که نورون‌ها در دسته‌هایی به نام لایه مرتب می‌شوند. معماری معمول ای‌ان‌ان متشکل از سه لایه است، لایه ورودی (داده‌ها را در شبکه توزیع می‌کند)، لایه پنهان (داده‌ها را پردازش می‌کند) و لایه خروجی (نتایج را به‌ازای ورودی‌های مشخص، استخراج می‌کند). یک شبکه می‌تواند چندین لایه پنهان داشته باشد. به‌هرحال، پژوهش‌های صورت‌گرفته در این زمینه نشان داده‌اند یک لایه پنهان برای این‌گونه مدل‌ها، هر تابع پیچیده و غیرخطی را تقریب می‌زند [۱۵]، همچنین، نتایج تجربی و عملی نیز این موضوع را تأیید می‌کنند [۴]. تقسیم‌بندی

1. Principal Components Analysis
2. Artificial Neural Network

بهترین نتیجه را بدهد به عنوان الگوریتم مناسب مسئله انتخاب کرد. پس از امتحان نه تابع آموزش و مقایسه نتایج آن‌ها مطابق شکل ۳ و جدول ۳، در نهایت الگوریتم لونبرگ-مارکوئیت (LM) بهترین کارایی را با توجه به حداقل زمان و کمترین خطای اجرا نشان داده و به عنوان الگوریتم آموزشی شبکه انتخاب شده است [۱۸].

می‌شوند. تعیین این مسئله که کدام الگوریتم برای یک مسئله خاص سریع‌تر عمل می‌کند، کار دشواری است، زیرا این مسئله به عامل‌های زیادی مانند پیچیدگی مسئله، تعداد داده‌های مجموعه آموزشی، تعداد وزن‌ها و بایاس‌ها در شبکه، مقدار خطای قابل قبول و در نهایت به اینکه شبکه برای تشخیص الگو یا تخمین تابع استفاده می‌شود، بستگی دارد. در نتیجه، باید توابع آموزشی مختلف را آزمود و تابعی که



شکل ۳. مقایسه الگوریتم‌های آموزشی مختلف

جدول ۳. مقایسه نتایج الگوریتم‌های آموزشی مختلف

الگوریتم	زمان متوسط (ثانیه)	نسبت	کمترین زمان (ثانیه)	بیشترین زمان (ثانیه)	میانگین مربع خطاها
LM	۲۲/۴۵	۱	۱۴/۰۲	۳۳/۰۳	۰/۳۴
BFG	۳۰/۰۸	۱/۳۴	۲۵/۷۸	۳۶/۷۸	۱/۰۵
SCG	۳۹/۹۶	۱/۷۸	۳۰/۷۰	۴۷/۸۷	۱/۷۸
CGF	۴۲/۶۶	۱/۹	۳۵/۶۰	۴۸/۹۰	۲/۵۹
CGP	۴۵/۳۵	۲/۰۲	۳۸/۶۵	۵۰/۴۳	۲/۸۳
CGB	۴۷/۱۵	۲/۱۰	۴۰/۵۷	۵۵/۳۴	۳/۰۵
OSS	۶۰/۶۲	۲/۷	۵۳/۲۴	۷۱/۰۷	۳/۵۳
RP	۸۰/۸۲	۳/۶۰	۷۸/۰۵	۸۹/۷۴	۷/۸۶
GDx	۱۶۷/۰۳	۷/۴۴	۱۴۳/۲۳	۱۸۰/۰۳	۱۳/۶۴

الگوریتم آموزشی شبکه و توابع فعال‌سازی تانژانت سیگمویت و تابع انتقال خطی پیورلین در لایه‌های پنهان و خروجی استفاده شده است. حداقل خطا، تعداد اپوک، گرادینان، مومنتوم و کنترل اعتبارسنجی به ترتیب دارای مقادیر ۰،۰۰۱، ۰،۰۰۰، ۱۰^{-۱}، ۱۰^{-۱۰} و ۴ هستند. همچنین، برای ارزیابی میزان

بنابراین، در این پژوهش از شبکه عصبی سه لایه با تعداد نورون مشخص در لایه میانی، تابع‌های انتقال تانژانت سیگمویت^۱ و پیورلین^۲ و الگوریتم لونبرگ-مارکوئیت به عنوان

1. Tansig
2. Purlin

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (۶)$$

که x_i و X_i به ترتیب داده‌های مشاهده‌ای و تخمینی هستند. همچنین:

R: ضریب همبستگی پیرسون،
RMSE: مجذور میانگین مربع خطاها و
MAE: میانگین مطلق خطا
هستند.

۴.۲.۲. تهیه مدل شبیه‌ساز
مراحل تهیه مدل شبیه‌ساز تصفیه‌خانه، در قالب فلوجارت شکل ۴ ارائه می‌شود [۱۸، ۳].

خطای مدل از سه مشخصه ضریب همبستگی پیرسون، مجذور میانگین مربع خطاها و میانگین مطلق خطا استفاده شده است که روابط ۲ تا ۶ معادله‌ها این معیارها را نشان می‌دهد [۱۹].

$$R = \frac{\sum_{i=1}^n (X_i - \bar{X})(x_i - \bar{x})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 (x_i - \bar{x})^2}} \quad (۲)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_i - x_i)^2}{n}} \quad (۳)$$

$$MAE = \frac{\sum_{i=1}^n |X_i - x_i|}{n} \quad (۴)$$

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad (۵)$$

جدول ۴. نتایج تحلیل نرم‌افزار SPSS برای مدل L. TDS

متغیرهایی که در مدل استفاده خواهند شد	ضریب KMO*	کد واحد		مدل
		۶۸۰۵	۶۸۱۸	
pH-COD-TDS-TSS-Turbidity-pH	۰/۷۰۵	pH		۱
بیش از ۹۰ درصد داده‌ها به صورت < ۵۰ ثبت شده بودند.	۰/۷۵۴	COD		۲
pH-COD-TDS-TSS-Turbidity-TDS	۰/۷۷۰	TDS		۳
تمام داده‌های ثبتی به صورت < 0.1 ثبت شده بودند.	-	Phenol	pH COD	۴
تمام داده‌های ثبتی به صورت < ۱۰ ثبت شده بودند.	-	TSS	TDS TSS	۵
pH-COD-TDS-Turbidity-Turbidity	۰/۶۷۴	Turbidity	Turbidity	۶
pH-COD-TDS-TSS-Turb.-Cl ⁻	۰/۶۸۴	Cl ⁻		۷
تمام داده‌های ثبتی به صورت < ۱۰ ثبت شده بودند.	-	Oil		۸

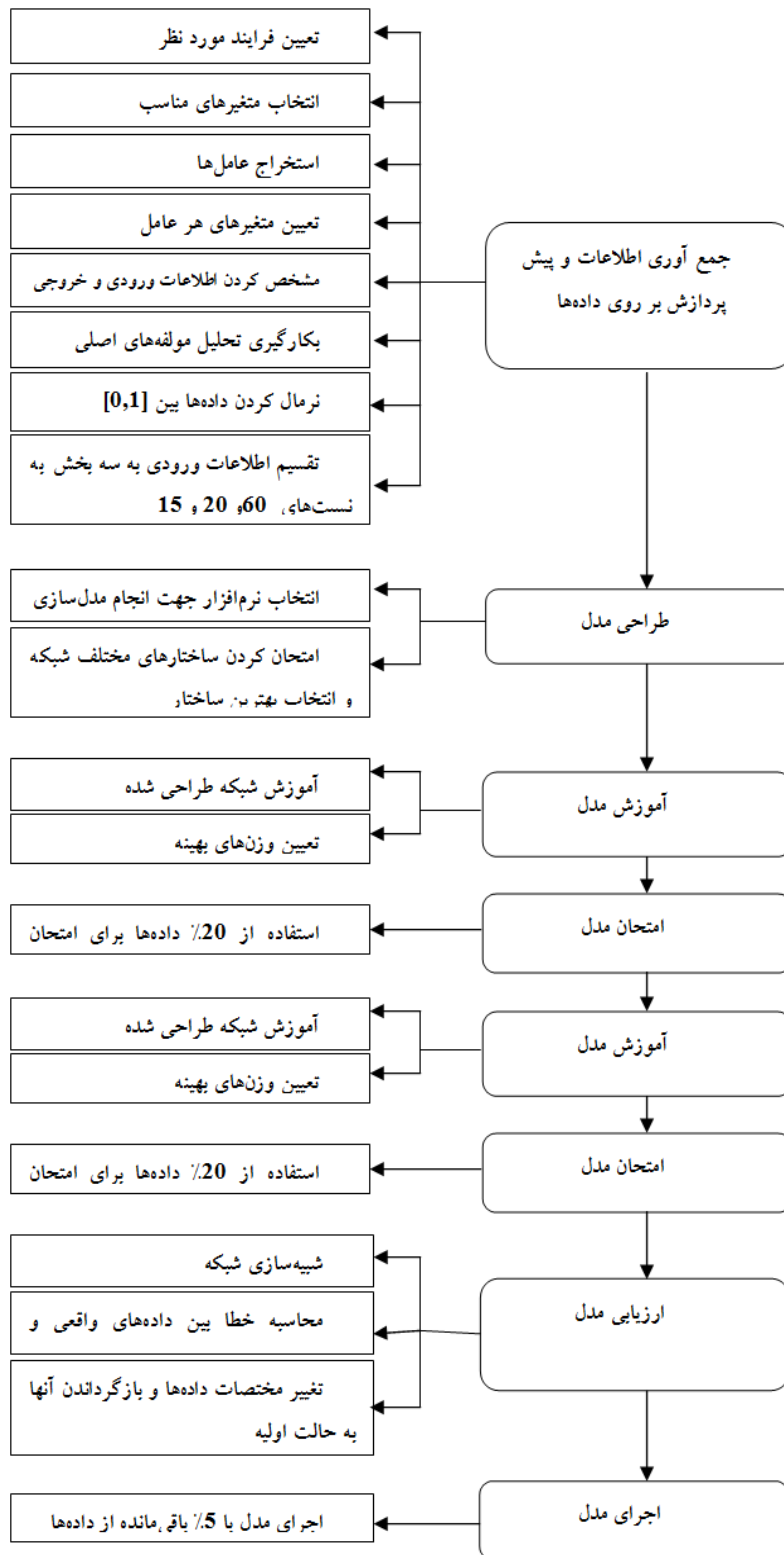
۳. نتایج و بحث

با توجه به توضیحات ارائه‌شده، در این قسمت با استفاده از اطلاعات در دسترس واحد تصفیه پساب با نمک پایین تصفیه‌خانه فجر، مدل‌سازی صورت گرفته است که در ادامه نتایج و تفسیر آن‌ها ارائه می‌شود. خروجی‌های مدل نیز در قالب نمودارها و جدول‌ها ارائه می‌شود که تمام آن‌ها نیز بررسی و تحلیل شده‌اند.

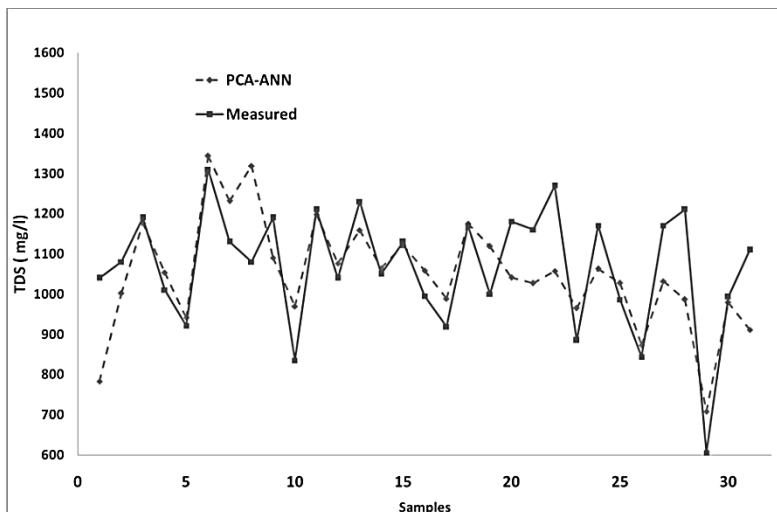
۱.۳. مدل L. TDS (واحد تصفیه پساب با نمک پایین)

در این قسمت، به منظور اطلاع از کیفیت پساب خروجی واحد مذکور، از مدل L. TDS استفاده شده است. به کارگیری این مدل پیش‌بینی، به راهبران تصفیه‌خانه امکان می‌دهد تا از میزان تقریبی کیفیت پساب خروجی در ابتدای ورود پساب

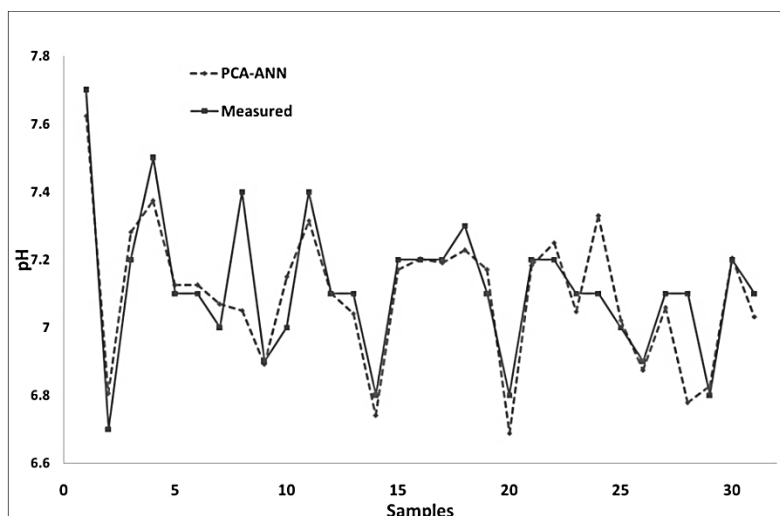
به تصفیه‌خانه اطلاع یابند. در نتیجه، استفاده از مدل پیش‌بینی، فرصت مورد نیاز را برای انجام‌دادن اقدامات لازم برای بهبود کیفیت در اختیار راهبران تصفیه‌خانه قرار می‌دهد. به منظور مدل‌سازی این واحد، ابتدا تحلیل عاملی برای شناسایی پارامترهای دارای قابلیت مدل‌سازی و دارای ارتباط معنی‌دار، در زمینه پارامترهای این واحد انجام گرفت. نتایج تحلیل در جدول ۴ آورده می‌شود. پس از مشخص شدن پارامترهای مناسب، مدل عصبی ترکیبی برای آن‌ها توسعه داده و اجرا شد. شکل‌های ۵، ۶، ۷ و ۸ نتایج مدل‌سازی این واحد را نشان می‌دهند. در پایان مدل‌سازی، خطای هر یک از مدل‌ها از طریق معیارهای خطا به دست می‌آید و قابلیت اطمینان مدل بررسی می‌شود. نتایج بررسی خطای مدل‌ها در جدول ۵ ارائه می‌شود.



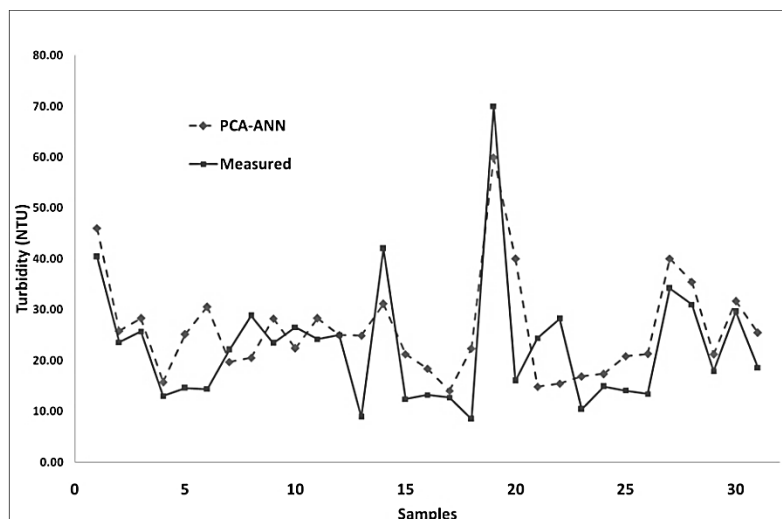
شکل ۴. الگوریتم ایجاد مدل شبیه‌ساز عملکرد تصفیه‌خانه [۱۸]



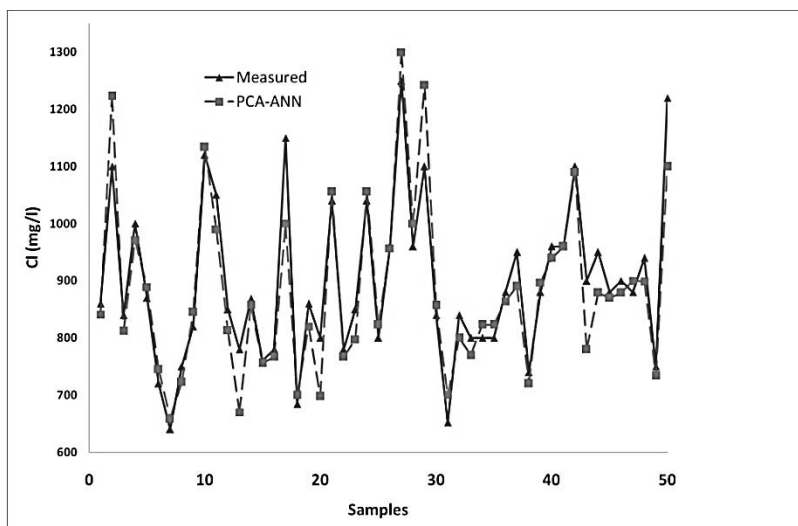
شکل ۵. خروجی مدل عصبی و نمودار مربوط به واحد TDS واحد TDS.



شکل ۶. خروجی مدل عصبی و نمودار مربوط به واحد pH واحد TDS.



شکل ۷. خروجی مدل عصبی و نمودار مربوط به کدورت واحد TDS.



شکل ۸. خروجی مدل عصبی و نمودار مربوط به Cl^- واحد TDS.

جدول ۵. نتایج مدل‌های عصبی واحد TDS.

ردیف	مدل	آموزش R	اعتبارسنجی R	تست R	کلی R	MAE	RMSE	تعداد داده‌ها
۱	pH	۰/۹۷	۰/۸۶	۰/۸۳	۰/۹۴	۰/۰۸	۰/۰۲	۴۱
۲	(NTU) Turbidity	۰/۸۸	۰/۷۸	۰/۸۶	۰/۸۳	۷/۱۷	۱/۵۹	۴۱
۳	(mg/l)TDS	۰/۹۰	۰/۸۲	۰/۸۱	۰/۸۴	۹۱/۵۴	۲۰/۹۷	۴۱
۴	(mg/l) Cl^-	۰/۹۱	۰/۶۶	۰/۷۸	۰/۸۰	۳۹/۴۲	۵۳/۹۸	۵۰

شکل ۷ نتایج مدل ترکیبی کدورت را نمایش می‌دهد. با توجه به نتایج تحلیل عاملی در هر واحد، مدل‌سازی پارامتر کدورت فقط در این واحد انجام گرفته است؛ بنابراین، امکان قیاس میزان تغییرات آن با مدل مشابه در واحدهای دیگر وجود ندارد. به‌هرحال براساس شکل ۷، بازه تغییرات این پارامتر بین حدود ۱۵ تا ۶۰ است و مدل توانایی لازم را به‌منظور پوشش مقادیر کمینه و بیشینه دارد.

براساس شکل ۸ مشاهده می‌شود مدل عصبی ترکیبی کلر تقریباً توانایی پوشش نوسانات اتفاق افتاده را در مقدار غلظت کلر دارد و رفتار تقریباً مشابه‌ای با داده‌های مشاهداتی دارد. همچنین، بازه تغییرات کلر بین حدود ۷۰۰ تا ۱۳۰۰ است.

جدول ۵ نتایج مدل‌های عصبی ترکیبی پارامترهای مدل‌سازی شدهٔ تانک ذخیره را ارائه می‌دهد. در تحلیل اولیهٔ جدول، مشاهده می‌شود مدل‌ها تقریباً میزان ضریب همبستگی بالایی دارند؛ بنابراین، انتظار می‌رود رفتار مشابه‌ای را با داده‌های مشاهداتی داشته باشند. این موضوع از طریق نمودارهای توضیح داده شده اثبات و مشاهده شد. همچنین، از میان مدل‌های این واحد، مدل پی‌اچ بیشترین ضریب

براساس جدول ۴، مشاهده می‌شود تحلیل عاملی برای پارامترهای دو واحد ۶۸۱۸ و ۶۸۰۵ انجام گرفته است. نتایج نشان می‌دهد پس از مقایسه و بررسی پنج پارامتر واحد خنثی‌سازی، با هشت پارامتر تانک ذخیره، فقط چهار پارامتر پی‌اچ، تی‌دی‌اس، کدورت و یون کلر، دارای همبستگی و ارتباط مناسب با ورودی‌ها هستند و قابلیت مدل‌شدن را دارند. در نتیجه، در این مرحله حدود ۲۰ درصد داده‌ها مدل‌سازی می‌شود.

شکل ۵ نمودار مدل عصبی ترکیبی تی‌دی‌اس را نمایش می‌دهد. طبق نمودار، بازه تغییرات تی‌دی‌اس نوسان دارد و این نوسان بین حدود ۷۰۰ تا ۱۳۵۰ میلی‌گرم بر لیتر است. مدل عصبی تی‌دی‌اس، توانایی پوشش‌دادن مقادیر کمینه و بیشینهٔ مشاهداتی را دارد و این موضوع با توجه به شکل ۵ اثبات می‌شود.

در بررسی نمودار پیش‌بینی مدل عصبی ترکیبی پی‌اچ (شکل ۶)، نکتهٔ شایان توجه شباهت زیاد نتایج پیش‌بینی و مقادیر مشاهداتی است. همان‌طور که مشاهده می‌شود، بازه تغییرات پی‌اچ در این واحد حدود ۶/۷ تا ۷/۷ بوده که مقادیر کمینه و بیشینهٔ آن به‌خوبی از طریق مدل پیش‌بینی پوشش داده شده است.

نتایج تحلیل عاملی نشان داد پارامترهای محدودی قابلیت مدل‌سازی دارند. براساس نتایج این تحلیل پارامترهای پی اچ، تی دی اس، کدورت و یون کلر مدل‌سازی شده‌اند. در ارزیابی میزان خطای مدل‌ها، نتایج نشان‌دهنده دقت تقریباً مناسب و قابل قبول مدل‌های پیش‌بینی است. همچنین، میزان ضریب همبستگی مدل‌های پیش‌بینی بالاتر از ۰/۸ است که با توجه به نوسان زیاد در غلظت پارامترهای ورودی، مقدار قابل قبولی است. نکته شایان ذکر دیگر، نبود مدل‌های مشابه و معیارهای لازم در شبیه‌سازی عصبی تصفیه‌خانه پساب صنعتی (پتروشیمی) است. براین اساس، امکان مقایسه مدل عصبی ترکیبی با مدل‌های مشابه وجود نداشته است و نتایج تحلیل انجام گرفته قابل پذیرش است. در نتیجه، نتایج این پژوهش به عنوان مبنایی برای مقایسه و بهبود مطالعات آتی در نظر گرفته می‌شود. از جمله پیشنهادها شایان ذکر برای ادامه مطالعات، استفاده از سایر روش‌های ترکیبی مدل‌سازی مانند مدل‌های عصبی- فازی و به‌ویژه مدل‌سازی عصبی- دینامیکی است.

همبستگی و دقت را دارد و مدل پیش‌بینی کلر دارای کمترین میزان همبستگی با داده‌های مشاهداتی است. در ارزیابی میزان خطای مدل کدورت، مشاهده می‌شود به علت کوچک بودن اعداد و کم بودن مقدار نوسان، مدل خطای تقریباً کمی داشته است. همچنین، در زمینه مدل پیش‌بینی کلر نیز می‌توان موضوع قابل قبول بودن را مطرح کرد. در نهایت، شایان ذکر است مدل پی اچ بیشترین قابلیت اطمینان و دقت مناسب را در پیش‌بینی دارد و بهره‌برداران می‌توانند از آن استفاده کنند.

۴. جمع‌بندی

براساس هدف پژوهش، با استفاده از روش ترکیبی شبکه عصبی و تکنیک تحلیل مؤلفه‌های اصلی برای مدل‌سازی بخش تصفیه پساب با نمک پایین تصفیه‌خانه فجر، مدلی ایجاد شد تا در کمترین زمان و با دقت مناسب، مشخصات کیفی پساب خروجی واحد را ارائه دهد. برای افزایش دقت مدل، از روش تحلیل عاملی به منظور تعیین داده‌های ورودی مؤثر به مدل استفاده شده است.

مراجع

- [1]. Mehrdadi, N., Hasanlou, H. (2012). "Investigating the Performance of Advanced Treatment Unit of Industrial Wastewater Treatment Plant Using Artificial Neural Network Model", International Conference on Chemical Processes and Environmental issues (ICCEEI'2012) Singapore.
- [2]. Homada, M. F., Al-Ghusian, I. A. (1999). "Integrated Wastewater Treatment Plant Performance Evaluation Using Artificial Neural Networks", Water Science and Technology, 40 (7), 55-66.
- [3]. Bowdena, G. J., Dandyb, G. C., Maier, H. R. (2005). "Input determination for neural network models in water resources applications", Journal of Hydrology, 301 (1-4), 75-92.
- [4]. Oliveira-Esquerre, K. P., Mori, M., Bruns, R. E. (2002). "Simulation of an Industrial Wastewater Treatment Plant Using Artificial Neural Networks and Principal Components Analysis", Brazilian Journal of Chemical Engineering, 19(4), 365-370.
- [5]. Kardam, A., Raj, K. R., Arora, J. K., Srivastava, M. M., Srivastava, S. (2010). "Artificial Neural Network Modeling for Sorption of Cadmium from Aqueous System by Shelled Moringa Oleifera Seed Powder as an Agricultural Waste", Journal of Water Resource and Protection, 2 (4), 339-344.
- [6]. Tezel, G., Yel, E., Sinan, R. K. (2010). "Artificial Neural Network (ANN) Model for Domestic Wastewater Treatment Plant Control", BALWOIS – Ohrid, Republic of Macedonia, 25- 29 may
- [7]. Elmolla, S. E., Chaudhuri, M. (2010). "The Use of Artificial Neural Network (ANN) for Modelling, Simulation and Prediction of Advanced Oxidation Process Performance in Recalcitrant Wastewater Treatment. ", Artificial Neural Networks – Application, InTech Open Access Publisher.
- [8]. Yetilmesoy, K., Sapci -Zengin, Z. (2009). "Stochastic Modeling Applications for the Prediction of COD Removal Efficiency of UASB Reactors Treating Diluted Real Cotton Textile Wastewater", Stochastic Environmental Research and Risk Assessment, 23(1), 13-26.
- [9]. Vyas, M., Modhera, B., Vyas V., Sharma, A. K. (2011). "Performance Forecasting of Common Effluent Treatment Plant Parameters by Artificial Neural Network", ARPN Journal of Engineering and Applied Sciences, 6 (1), 38-42.

- [10]. Hasanlou, H., Mehrdadi, N., Jafarzadeh, M. T., Hasanlou, H. (2012). "Performance Simulation of H-TDS Unit of FAJR Industrial Wastewater Treatment Plant Using a Combination of Neural Network and Principal Component Analysis", *Journal of Water Resource and Protection*, 4 (5), 311-318.
- [11]. Mehrdadi, N., Hasanlou, H., Jafarzadeh, M. T., Hasanlou, H., Abdolabadi, H. (2012). "Simulation of Low TDS and Biological Units of Fajr Industrial Wastewater Treatment Plant Using Artificial Neural Network and Principal Component Analysis Hybrid Method", *Journal of Water Resource and Protection*, 4 (5), 370-377.
- [12]. Zhao, H., Hao, O., McAvoy, T., and Chang, C (1997). "Modeling nutrient dynamics in sequencing batch reactor", *Journal of Environmental Engineering*, 123 (4), 311-319.
- [13]. Shariat zadeh, M. (2010). "Fajr Petrochemical Environmental Landscape", National Petrochemical of IRI.
- [14]. Zare Chahooki, M. (2011), "Methods of Multivariate Analysis in SPSS", Department of Natural Resources, University of tehran, Iran.
- [15]. Hornik, K., Stinchcombe, M., and White, H. (1989). "Multilayer feedforward networks are universal approximators", *Neural Networks*, 2(5), 359-366
- [16]. Stone, M. (1974). "Cross-validators Choice and Assessment of Statistical Predictions", *Journal of the Royal Statistical Society*, 36 (2), 111-147.
- [17]. Burden, F. R., Brereton, R. G., Walsh, P. T. (1997). "Cross-validators selection of test and validation sets in multivariate calibration and neural networks as applied to spectroscopy", *Analyst*, 122 (10), 1015-1022.
- [18]. Hasanlou, H. (2012). "Simulation Of Industrial Waste Water Treatment Plant Using Artificial Neural Network And Principal Component Analysis Hybrid Method", MS. Thesis, University of Tehran, Iran.
- [19]. Memarian, H., Balasundram, S. K. (2012). "Comparison between Multi-Layer Perceptron and Radial Basis Function Networks for Sediment Load Estimation in a Tropical Watershed", *Journal of Water Resource and Protection*, 4(10), 870-876.