نشریه علمی– پژوهشی امیرکبیر (مهندسی عمران و محیط زیست)
Amirkabir Journal of Science & Research (Civil & Environmental Engineering)
(AJSR - CEE)

دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)
دوره چهل و شش، شماره۲، زمستان ۱۳۹۳، صفحه ۲۹ تا ۳۷
Vol. 46, No. 2, Winter 2014, pp. 29- 37

بررسی میزان تأثیر پیش پردازش داده ها در دقت نتایج مدلسازی تولید پسماند شهری با استفاده از شبکه عصبی

ملیحه فلاح نژاد[۱]، محمد عبدلی[۲*]

۱- دکتری، دانشکده محیط زیست، دانشگاه تهران
۲- استاد، دانشکده محیط زیست، دانشگاه تهران

(دریافت۱۳۹۰/۳/۲۳، پذیرش۱۳۹۳/۱۰/۲۹)

چکیده

تولید پسماند در جوامع بشری امری روزمره و طبیعی است. پسماند از مرحله تولید تا مرحله مصرف و مرحله دفع نهایی تولید شده و امری غیر قابل اجتناب است. توسعه شهرها و صنعتی شدن آنها باعث تولید روزافزون پسماند شهری می‌شوند. برای آگاهی از کمیت این پسماندها گامی ضروری است. در این پژوهش از شبکه عصبی مصنوعی به عنوان یک ابزار کارآمد برای مدل‌سازی میزان پسماند تولیدی شهر مشهد استفاده شده‌است. در این راستا ابتدا پیش‌پردازش‌هایی بر روی داده‌های ورودی متغیرهای مستقل و وابسته انجام شده که اثر اعمال هر روش بر روی دقت مدل تخمین زده شده بررسی می‌شود. نتایج نشان می‌دهد که با انجام پیش پردازشهایی بر روی داده های خام ورودی به مدل، می توان نتایج دقیقتری بدست آورد. سه حالت مختلف بررسی شد و بهترین پیش پردازش شامل لگاریتم‌گیری، حذف روند و استانداردسازی داده ها انتخاب شد. معماری شبکه دو لایه مخفی هرکدام با ۵ نرون و با MAPE ۰/۰۶، MSE ۰/۴۶ و ضریب همبستگی ۰/۸۶ بدست آمد.

کلمات کلیدی

پسماند شهری، شبکه عصبی، پیش پردازش، مشهد.

* نویسنده مسئول وعهده دار مکاتبات Email: mabdoli@ut.ac.ir

۱- مقدمه

پسماند شهری نتیجه طبیعی فعالیت های انسان است. مساله جمع‌آوری، حمل و نقل و دفع مواد زائد جامد شهری در جهان یکی از مشکلات و معضلات شهری به شمار می‌رود. تولید انبوه زباله در شهرها معضلات بهداشتی، زیست‌محیطی، اقتصادی- اجتماعی، حمل و نقل و ترافیک را به وجود آورده است. مکان گزینی محل دفن و صنایع بازیافت مواد زائد یکی از مسائل عمده و پیچیده در شهرهاست. در صورتی که سامانه مدیریتی مناسبی به کار گرفته نشود، این مواد باعث آلودگی‌های زیست محیطی زیادی می‌شوند و سلامت بشر را به خطر می اندازند [۱]. ایجاد چنین سیستمی به دلیل پیچیدگی و طبیعت بسیار ناهمگون پسماند، بسیار مشکل می باشد. یکی از عوامل بسیار مهم در راه‌اندازی صحیح چنین سیستم پیچیده‌ای شناخت کمیت زباله تولیدی می- باشد، زیرا کمیت تولید در حجم سرمایه‌گذاری برای ماشین‌آلات، مخازن ذخیره در محل، ایستگاه‌های انتقال، ظرفیت دفع، سازماندهی و تشکیلات مناسب و تعیین محدوده توسعه، مؤثر است. همچنین کمبود اطلاعات و آمار دقیق، نبود برنامه‌ای جامع در مدیریت پسماند شهری، نبود ثبات قیمت‌ها و امکان بروز بحران‌های محتمل پیش‌رو به جامع‌نگری و برخورد سیستمی در مدیریت پسماند شهری اهمیت ویژه‌ای بخشیده است.

تخمین میزان پسماند تولیدی به دلیل نوسانات زیادتولید و عوامل گوناگونی که برآن مؤثر است، یکی ازکارهای بسیار دشوار در امر مدیریت پسماند می‌باشد. از آنجا که طراحی سیستم‌های مدیریت پسماند شهری باید برای یک بازه طولانی مدت انجام شود، لذا آگاهی از کمیت پسماند در افق طرح ضروری است. تعیین میزان پسماند تولیدی در دراز مدت با استفاده از روش‌های قطعی و غیر قطعی می تواند اطلاعات مناسبی در امر طراحی و برنامه‌ریزی سیستم مدیریت پسماند شهری در اختیار مدیران و طراحان قرار دهد.

در ایران در سیستم مدیریت مواد زائد، سالانه بیش از ۲۲۰ میلیارد تومان هزینه می شود، با این حال هنوز نمی توان به استانداردهای بین المللی حتی نزدیک شد. در سال ۱۳۸۳ پسماند شهر مشهد به ۱۷۰۰ تن و خاک و نخاله آن به حدود۶۰۰۰ تن رسیده است . نبود برنامه ریزی دقیق موجب شده است که صنایع و کارگاههای کوچک غیر استاندارد بازیافت در حواشی شهرها یعنی در مکانهایی که نیروی کار ارزان و به وفور یافت میشود و کنترلهای زیست‌محیطی و اجتماعی به حداقل می‌رسد، مستقر شوند. لذا به منظور اتخاذ تدابیر لازم در امر پردازش و بازیافت، کمپوست و دفع نهایی آگاهی از کمیت پسماند تولیدی در درجه اول امری ضروری است. در سالهای اخیر مطالعات زیادی در مورد تخمین و مدل سازی میزان زباله تولیدی انجام شده است.

داسکالوپوس و همکارانش[۱] برای اروپا و آمریکا، اثر متغیرهای GDP و جمعیت را بر روی تولید پسماند بررسی نمودند. برای GDPاز بازه زمانی ۱۹۸۰-۱۹۹۳ برای اروپا و ۱۹۶۰- ۱۹۹۳برای آمریکا استفاده نمود. یک معادله غیرخطی توانی که GDP و جمعیت بعنوان متغیر مستقل و پسماند بعنوان متغیر وابسته بدست آمد. سپس میزان تولید هر بخش از پسماند مانند کاغذ و غیره را با معادله ای که تابع کل هزینه های مصرف‌کننده بود بدست آورد [۱].

بیگل و همکارانش[۲] با سناریوهای مختلف و با استفاده از مدل رگرسیونی و روش حداقل مربعات اثر متغیرهای مستقل GDP ،مرگ و میر نوزادان، امید به زندگی و جمعیت را بر روی پسماند تولیدی برای شهرهای اروپا بدست آوردند. R2 مدل حاصله ۶۵٪ بدست آمد و درصد خطای مطلق ۸% بدست آمد. این میزان R2 و خطا نشان دهنده آنست که مدل تخمین زده شده تا حد مطلوبی می تواند میزان تولید پسماند را توصیف نماید [۲].

دیسون و همکارانش[۳] از روش سیستم های دینامیکی و نرم افزار استلا[۴] برای پیش بینی پسماند تولیدی در مناطق شهری با رشد زیاد استفاده نمودند. برای تحقق مدل از داده های شهر سن آنتونیو در تگزاس استفاده شد. مدلسازی با شرکت متغیرهای رشد جمعیت، درآمد خانوار، فعالیت اقتصادی و تعداد ساکنین در هر خانه انجام شد. اطلاعات مربوط به این متغیرها تنها برای سالهای ۱۹۸۰، ۱۹۹۰، ۲۰۰۰ تهیه شد و بدلیل کمبود اطلاعات سیستم دینامیکی یک روش مفید شناخته شد و پیش بینی تا سال ۲۰۱۰ انجام شد. R2 مدلها بین ۰/۸۹ و ۰/۹۹ بدست آمد [۷].

سوکو و همکارانش[۵] از میزان پسماند تولیدی و ترکیبات آن در فنلاند بین سالهای ۱۹۶۰ و ۲۰۰۲ استفاده کرد نمودند و اثر میزان جمعیت، سطح رفاه و سطح فناوری را بر روی پسماند تولیدی تحلیل کرد. آنها در تحقیقاتشان از معادله IPAT استفاده کردند که در آن I اثر زیست محیطی، p جمعیت، A سطح رفاه و T سطح فناوری است. این معادله در مطالعات زیادی مورد استفاده قرار گرفت. با در نظر گرفتن متغیرهای مختلف میزان تولید پسماند در فنلاند را تا سال ۲۰۲۰ پیش بینی کردند. در این تحقیق از GDP بعنوان شاخص رفاه استفاده شد و با در نظر گرفتن درصد تغییرات هر یک از متغیرها با سه سناریو، هر متغیر تا ۲۰۲۰ پیش بینی شد. بررسی داده های گذشته در فنلاند نشان داد که میزان پسماند تولیدی در فنلاند تا سال ۱۹۹۰ افزایش یافته، سپس تا سال ۱۹۹۷ کاهش یافته و بعد تا ۲۰۰۰ افزایش و پس از آن کاهش می یابد. میزان مواد آلی و پلاستیک در دوره مطالعه افزایش و میزان کاغذ و مقوا کاهش یافته است [۶].

سارا اجدا بنیتز و همکارانش[۶] سه مرحله آمارگیری در مورد کمیت و کیفیت پسماند تولیدی از چند محله در سه سال مختلف

انجام داد. متغیرهای مستقل برای این مناطق، تعداد ساکنین در هر خانه، درآمد خانوار، سطح تحصیلات و متغیر وابسته، پسماند سرانه تولیدی بود. او معادله رگرسیونی خطی یک متغیره، دو متغیره و سه متغیره تخمین زد و بهترین معادله که دارای R2،۰/۵۱ بود و دارای متغیرهای مستقل درآمد خانوار، سطح تحصیلات و تعداد ساکنین در هر خانه را انتخاب کرد [۹].

بچ و همکاران[۷] مدلی برای پیش بینی مقدار پسماند کاغذی تخمین زدند. این مدل شامل ۸ متغیر مستقل: تعداد ایستگاه های شبانه به ازای هر نفر، شاخص قدرت خرید برای هر خانوار، شاخص قدرت خرید برای هر نفر، تعداد خانم های خانه‌دار، تعداد کارمندان در ادارات و صنعت، تعداد کارمندان در بخش خدمات، تعداد شرکتهای کشاورزی و تعداد مکانهای جمع آوری کاغذ می شود. R^2 مدل تخمین زده شده ۴۸٪ بدست آمد [۳].

نوری و همکارانش دو روش پیش بینی را برای ارزیابی میزان پسماند تولیدی در شهر تهران مقایسه نمودند. بدین منظور داده های هفتگی پسماند تولیدی شهر تهران را با ۱۲ تأخیر فاز (برابر با یک فصل) استفاده کرد. در این پژوهش روش رگرسیون چند متغیره خطی MLR بهمراه PCA (روشی برای حذف هم خطی) با ANN مقایسه شد. در روش شبکه عصبی، از روش MLP استفاده شد. ضریب همبستگی R و AARE در شبکه عصبی به ترتیب برابر با ۰/۸۳ و ۴/۴٪ بدست آمد. در مقایسه با PCA-MLR که ضریب همبستگی ۰/۴۴۵ و MAPE ۶/۶٪ بدست آمده است، مدل شبکه عصبی نتایج بهتری دارد. بنابراین مدل شبکه عصبی برای پیش بینی پسماند تولیدی شهر تهران به صورت کوتاه مدت نتایج بهتری می دهد [۱۰].

نوری و همکارانش برای بدست آوردن مدل پیش بینی پسماند شهری مشهد به صورت هفتگی از روش‌های (SupportVector Machine (SVM و PCA در شبکه عصبی استفاده کرد. در این تحقیق با استفاده از تکنیک PCA، روی داده ها پیش پردازش انجام شده وسپس این داده ها وارد مدل SVM شده است. در این مدل داده‌های ورودی پسماند شهر مشهد به صورت هفتگی از سال ۲۰۰۵ تا ۲۰۰۸ استفاده شد، و نتایج پیش‌بینی با استفاده از مدل SVM و داده‌های واقعی مقایسه شد. از آنجایی که در این مدل نیز تنها از دادهای پسماند بصورت خود همبسته استفاده شده است بنابراین تنها پیش‌بینی برای یک یا دو مرحله (در اینجا یک یا دو هفته) قابل استناد است و برای پیش بینی هفته های بعدی به دلیل افزایش خطا در پیش‌بینی (چون برای پیش‌بینی مراحل بعدی از پیش‌بینی گام قبلی استفاده می‌نماید که این کار خطا را بسیار افزایش می دهد) نمی‌توان به مدل تخمین زده شده استناد نمود [۱۱].

چونگ و همکاران[۸] ابتدا مزایا و معایب مدل‌های سری زمانی و چند عاملی را بررسی نموده و سپس با استفاده از مدل تحلیل چند فاکتوری، میزان پسماند تولیدی را تا سال ۲۰۳۶ پیش بینی نمودند. در تحقیق آنها، GDP سرانه، جمعیت، و تعداد مساکن بعنوان متغیرهای مستقل در نظر گرفته شده و میزان پسماند تولیدی روزانه سرانه هم بعنوان متغیر وابسته در نظر گرفته شده است و بهترین مدل اتورگرسیونی تخمین زده شد. سپس با پیش بینی متغیرهای ورودی میزان پسماند تولیدی صنعتی و خانگی را تا سال ۲۰۳۶ پیش بینی کرد [۴].

در تحقیقات گذشته اغلب از شبکه عصبی برای پیش بینی کوتاه مدت استفاده شده است و از این روش‌ها فقط می‌توان تا چند گام زمانی مقادیر را تولید کرد. همچنین از مدل‌های خود همبسته استفاده شده است که با اطلاعات پسماند گذشته، آینده پیش‌بینی می‌شود، این رویکرد باعث افزایش خطا در طول بازه پیش‌بینی خواهد شد. لذا در این مقاله راهکاری برای رفع مشکل شبکه عصبی در پیش‌بینی بلند مدت ارائه شده و نتایج نشان می-دهند با انجام این راهکار دقت مدل شبکه عصبی تا حد زیادی بهبود می‌یابد.

با توجه به اهمیت آگاهی از کمیت پسماند تولیدی در سال‌های آینده به منظور طراحی سیستم مدیریت مواد زائد و انجام برنامه-ریزی‌های لازم برای پردازش، بازیافت و دفع، در این تحقیق با در نظر گرفتن متغیرهای اقتصادی اجتماعی و آب و هوایی مختلف امکان سنجی پیش‌بینی بلند مدت تولید پسماند در سال‌های آینده (۲۲ سال) با استفاده از نرم‌افزار ۲۰۰۸ متلب۹ و شبکه عصبی مصنوعی صورت خواهد گرفت.

۲- مواد و روشها

هدف این تحقیق ارائه یک روش مناسب و قابل اطمینان برای بدست آوردن میزان تولید بلند مدت پسماند شهری است، تا تصمیم‌گیری برای آینده با عدم قطعیت کمی صورت پذیرد. بدین منظور متغیرهای تاثیرگذار بر تولید پسماند شهری همچون متغیرهای اقتصادی اجتماعی و اقلیمی باید در نظر گرفته شوند و از این طریق میزان تولید پسماند شهری در آینده محاسبه می‌شود.

درگذشته با توجه به عدم محدودیت زمین، کوچک بودن شهرها، نبود سیستم‌های حمل و نقل پیشرفته پسماند شهری اغلب بصورت تلنبار زباله در محدوده شهرها دفع می‌شد و دفع مواد زائد مشکل خاصی را ایجاد نمی‌کرد. امروزه به دلیل تراکم جمعیت بسیار بالا، گسترش شهرها، گران بودن زمین، گران بودن ماشین‌آلات حمل و نقل و نیروی کارگر، نبود سوخت کافی برای حمل و نقل توسط ماشین‌آلات، مسئله مدیریت مواد زائد تولیدی

از تولید تا دفع یک مسئله حیاتی در جوامع در حال توسعه است. همین امر باعث می‌شود تا مدیران سیستم‌های مدیریت مواد زائد، نیازمند اطلاع دقیق از میزان پسماند تولیدی و جمعیت تحت پوشش سیستم جمع‌آوری پسماند باشند تا آینده‌نگری را بر اساس آن انجام دهند. نکته‌ای که در این بخش قابل توجه است، آن است که بتوان بر اساس تخمین میزان پسماند تولیدی شهری دقیق از یک منطقه، در مورد افزایش ظرفیت زمین‌های موردنیاز برای دفن بهداشتی، افزایش ظرفیت سیستم انتقال، تعداد نیروی کارگر، تعداد ماشین‌آلات و میزان حجم سرمایه گذاری تصمیم‌گیری کرد. بطور کلی پیش‌بینی میزان پسماند تولیدی به گونه‌ای که سیاست-گذاری‌های مدیریت مواد زائد مطابق واقعیت موجود اتخاذ شود، از طریق داشتن یک مدل دقیق و جامع پیش بینی پسماند تولیدی امکان پذیر است.

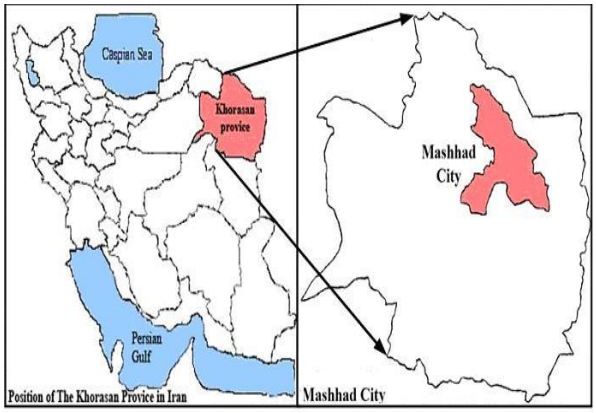
با توجه به بررسی و مطالعات انجام شده دیده شد، شبکه عصبی توانسته جواب‌های بهتری را نسبت به مدل‌های پیش‌بینی دیگر برای پدیده‌های غیر خطی ارائه نماید. از سوی دیگر در آن مطالعات برای پیش‌بینی تولید پسماند شهری، آن را تابعی از میزان پسماند تولیدی در سال‌های قبل در نظر گرفته اند، به همین دلیل و همچنین عدم قطعیت بالا در پیش‌بینی، تا چند مرحله زمانی قادر به پیش بینی تولید پسماند در آینده نیستند.

بنابراین، در این تحقیق برای بدست آوردن رابطه‌ی بین پارامترهای مستقل و وابسته و پیش‌بینی پسماند شهری تولیدی از شبکه عصبی مصنوعی طوری استفاده می‌شود تا بتوان تولید پسماند را بطور بلند مدت پیش‌بینی کرد. بدین منظور قبل از استفاده از داده‌ها، روی آن ها پیش پردازش انجا می‌شود. بعد از تعیین مدل شبکه عصبی بهینه، پارامترهای مستقل پیش‌بینی شده و به عنوان ورودی به شبکه عصبی داده می‌شود تا تولید پسماند در آینده حاصل شود. بنابراین با بدست آمدن پیش‌بینی برای دما، جمعیت و سایر متغیرهای موثر بر تولید پسماند می‌توان اثر تغییر آنها و بحران‌های محتمل در سال‌های آتی در محدوده مورد مطالعه را بر روی تولید پسماند شهری مورد بررسی قرار داد تا به این وسیله با اطلاع از وضعیت آتی در محدوده مورد مطالعه، با برنامه‌ریزی و مدیریت، سیاست مناسبی در مورد افزایش زمین-های مورد نیاز دفن بهداشتی، ارتقاء سیستمهای بازیافت و تبدیل مواد، میزان حجم سرمایه گذاری، تعداد ماشین آلات حمل و نقل و غیره اتخاذ نمود.

برای بررسی امکان‌سنجی پیش بینی بلندمدت تولید پسماند با شبکه عصبی، سه حالت مختلف بررسی میشود؛ ابتدا داده های سری زمانی متغیرهای ورودی و خروجی بدون هیچ پیش پردازشی در برنامه شبکه عصبی قرار می‌گیرند و نتایج بررسی می‌شوند. در حالت بعد، روند موجود در سری زمانی داده‌های ورودی و خروجی حذف می‌شود و با این حالت نیز مدلسازی انجام می‌شود. در حالت سوم ابتدا از داده‌های سری‌های زمانی ورودی و خروجی لگاریتم گرفته می شود و پس از لگاریتم‌گیری روند سری‌های زمانی حذف می‌شود. در این تحقیق به بررسی این سه حالت پرداخته شده و پارامترهای عملکرد بهینه در سه حالت با هم مقایسه خواهند شد.

۳- مطالعه موردی

مشهد شهری در شمال شرقی ایران و مرکز استان خراسان رضوی است. شهر مشهد مرکز استان خراسان رضوی با ۲۰۴ کیلومتر مربع مساحت می باشد. آب و هوای مشهد معتدل و متغیر است و وزش باد در آن بیشتر در جهت جنوب شرقی به شمال غربی است. حداکثر درجه حرارت در تابستان ۳۵ درجه بالای صفر و کمترین آن در زمستان ۱۵ درجه زیر صفر می‌باشد. بر اساس سرشماری عمومی نفوس و مسکن سال ۱۳۸۵، این شهر با ۲۴۱۰۸۰۰ نفر جمعیت دومین شهر پرجمعیت ایران پس از تهران است. در شکل (۱) موقعیت شهر مشهد در استان خراسان رضوی و در ایران نشان داده شده است.



شکل (۱): موقعیت شهر مشهد در کشور ایران

۴- شبکه عصبی

شبکه های عصبی روش‌های ریاضی ساده‌ای هستند که برای اجرای کارهای مختلف طراحی شده‌اند. در دهه ۱۹۷۰ میزان علاقه در این زمینه کم شد اما بعد از آن در دهه ۱۹۸۰ رشد بسیار زیادی در زمینه شبکه عصبی صورت گرفت. امروزه، ANN ها می توانند برای اجرای وظایف مختلفی مانند پیش‌بینی و مدل‌سازی استفاده شوند. ANN ها شامل ویژگی‌هایی هستند که راه حل‌های کاملی در سیستم‌های پیچیده‌ی خطی و غیرخطی ارائه می‌دهند. برخی از این ویژگی‌ها عبارتند از: قابلیت یادگیری، تعمیم، پردازش موازی و غیره. این ویژگی‌ها باعث می‌شوند ANN ها مشکلات پیچیده را به طور دقیق و قابل انعطاف حل نمایند. با توجه به اینکه تولید پسماند، ساختاری غیرخطی دارد ANN ایده‌ی خوبی

برای پیش‌بینی در این زمینه است.

در سال‌های اخیر، تحقیقات زیادی درمورد کاربرد شبکه‌های مصنوعی برای پیش‌بینی کوتاه مدت (هفته‌ای یا روزانه) یا میان مدت انجام شده‌است. چون شبکه‌های عصبی براساس محدوده‌ی اطلاعات اولیه (ورودی و خروجی) آموزش می‌یابند نمی‌توانند خارج از این محدوده آموزش دیده را پیش بینی کنند. در اکثر مطالعات قبلی در زمینه پیش‌بینی با استفاده از شبکه عصبی، از مدل‌هایی استفاده شده است که متغیر وابسته (تولید پسماند) به عنوان متغیر مستقل با وقفه زمانی در مدل آورده شده است. این امر باعث می شود برای ساخت مقادیر آینده پسماند تولیدی، از مقادیر پسماند (پیش بینی شده در مراحل قبل) استفاده شود که در مراحل قبل توسط همین مدل بدست آمده است. بنابراین خطا در هر مرحله افزایش یافته و این مدل‌ها قادر به ساخت اطلاعات تا چند مرحله زمانی هستند [۸]. در مدل‌های چند متغیره که تولید پسماند بصورت تاخیری از زمان بعنوان متغیر مستقل در مدل وجود دارد، بدلیل همبستگی بسیار زیاد بین متغیر وابسته و وقفه‌های همین متغیر که بعنوان مستقل در مدل قرار دارد، این مقادیر تأخیری بیشترین تأثیر را در پیش‌بینی تولید پسماند می-گذارد. بهمین دلیل اثر سایر متغیرها در پیش‌بینی مقدار تولید پسماند ناچیز است و نمی‌تواند اثر سایر عوامل موثر را بخوبی مدل کند.

می‌توان بصورت خلاصه مشکلات در پیش بینی بلندمدت با شبکه عصبی را بصورت زیر بیان کرد:

۵- درون یاب بودن شبکه های عصبی

افزایش خطا در پیش‌بینی ناشی از استفاده از مقادیر تاخیری متغیر وابسته بعنوان متغیر مستقل.

در نظر نگرفتن اثر سایر متغیرهای موثر بر تولید پسماند، بدلیل همبستگی زیاد بین تولید پسماند آینده و مقادیر تولید در وقفه-های قبلی.

اگرچه ANN تابحال برای پیش‌بینی کوتاه مدت استفاده شده است، در این قسمت نشان داده شده است که یک راهکار دقیق برای پیش‌بینی بلند مدت تولید پسماند است.

در این تحقیق، برای اینکه اثر متغیرهای موثر تولید پسماند دیده شود، تابع تولید پسماند شهری برای شهر مشهد با استفاده از شبکه عصبی تخمین زده شده است. از مدل شبکه عصبی پرسپترون چند لایه با الگوریتم پس انتشار خطا برای بدست آوردن رابطه بین متغیرهای مستقل و وابسته، برای پیش‌بینی پسماند تولیدی استفاده شده است.

بنابراین در این قسمت از متغیرهای تاخیری در مدل که موجب کاهش در دقت پیش‌بینی می شود مانند تحقیقاتی که تولید پسماند را پیش بینی نموده اند، استفاده نشده است زیرا در پیش‌بینی بلند مدت باعث افزایش خطا در پیش بینی می شود.

در این مدل از متغیرهای تابع (۱) استفاده شده است:

$$Sw = f(I, P, MT) \quad (۱)$$

که در آن:

Sw: مقدار تولید پسماند شهری (تن)

I: درآمد سرانه خانوار(ریال)

P: جمعیت(نفر)

MT: متوسط دمای حداکثر

با توجه به این امر که شبکه پرسپترون چند لایه (MLP) معمولا برای حل مسائل مهندسی بخوبی جواب می‌دهد، این روش انتخاب شده است. برای آموزش شبکه عصبی از روش پس انتشار خطا استفاده شده است که در این روش ورودی، خروجی شبکه وخروجی مطلوب برای آموزش شبکه به کار می‌روند تا شبکه دارای خروجی نزدیک به خروجی مطلوب باشد. تابع فعالیت مورد استفاده در شبکه BPNN از نوع تابع زیگموئیدی تانژانت هایپربولیک می‌باشد. خروجی این تابع همواره در بازه (۱:۱-) قرار دارد.

برای اجرای شبکه عصبی در این تحقیق از برنامه‌نویسی در نرم‌افزار MATLAB 2008 استفاده شده است و در آن به ترتیب برای دوره آموزش، صحت‌سنجی و آزمون از حدود ۷۰، ۱۵ و ۱۵ درصد از داده‌های موجود استفاده شده است. برای انتخاب بهترین ساختار شبکه عصبی، با تغییر در تعداد لایه های میانی (بین ۱ تا ۲ لایه)، اندازه لایه‌های میانی (بین ۴ تا ۱۰ نرون) و تغییر در انواع الگوریتم‌های موجود در متلب، بهترین ساختار با استفاده از چند حلقه از بین ۱۱۷۶ شبکه عصبی انتخاب شده است.

۶- پیش پردازش داده‌ها

برای اینکه بتوان مشکل درون گرایی شبکه عصبی را حل نمود و از آن در پیش‌بینی بلند مدت استفاده کرد لازم است تا روی داده ها تحلیل شده و تغییراتی صورت بگیرد. بعد از آموزش شبکه عصبی با داده‌های ورودی و خروجی برای بازه‌ی (۱۳۷۹-۱۳۸۸)، لازم است تا مقادیر ورودی برای دوره‌ی آتی ۲۰ ساله پیش‌بینی شود و بعنوان ورودی به مدل شبکه عصبی داده شود تا مقادیر تولید پسماند برای آینده بدست آید. با توجه به اینکه متغیرهای مستقل در طول زمان تغییرات زیادی نسبت به اطلاعات مشاهداتی داشته است و مدل در محدوده اطلاعات مشاهداتی آموزش دیده است بنابراین مدل قادر نیست این مقادیر را ببیند و مقادیر

خروجی متناظر با این متغیرهای مستقل را بدست آورد.

برای رفع این مشکل سعی شده است داده‌های ورودی و خروجی در دوره آموزش با داده‌های ورودی در دوره‌ی پیش‌بینی (۱۴۱۰-۱۳۹۰) هم مقیاس شوند. بدین منظور از مفهوم ایستایی استفاده شده است.

یک متغیر سری زمانی هنگامی ماناست (ایستا) که میانگین، واریانس و ضرایب خود همبستگی آن در طول زمان ثابت باقی بماند [۱۴]. سری‌های زمانی را می‌توان به اجزای روند، فصلی و جزء نامنظم تجزیه نمود. جزء روند، میانگین سری داده ها را در طول زمان تغییر می‌دهد و قسمت فصلی تغییرات دوره‌ای سری زمانی است که در یک بازه‌ی زمانی مشخص بطور متناوب این تغییرات تکرار می‌شود. جزء تصادفی ناشی از طبیعت غیر قطعی یک سری زمانی است. در مدل سازی آماری مهمترین مرحله مدل سازی بخش تصادفی سری زمانی است.

بنابراین پیش‌بینی بر اساس جزء تصادفی که در اغلب مدل‌های سری زمانی از این جزء استفاده شده است علاوه بر اینکه باعث افزایش دقت مدل سازی می‌شود، باعث می‌شود تا داده‌ها هم-مقیاس‌تر شوند. به منظور آماده‌سازی سری اطلاعات برای مدل سازی با شبکه عصبی، سعی شده است تا ترم های قطعی حذف شوند و داده‌ها به حالت ایستایی نزدیکتر شوند. بدین منظور چندین حالت بررسی شده است. در حالت اول از داده‌ها بدون هیچ‌گونه تغییری استفاده شده و فقط نرمال می‌شوند و در بازه -۱ و ۱ قرار می‌گیرند. در حالت دوم سعی شده است تا روند موجود کلیه داده‌ها حذف شود و سپس داده‌ها نرمال شوند. نرمال کردن داده‌ها بدین منظور است که از اشباع شدن نرون‌ها جلوگیری شود، زیرا اگر نرون‌ها اشباع شوند، تغییرات در ورودی تغییر کمی در خروجی ایجاد می‌نماید [۱۳]. برای رسیدن به ایستایی سری زمانی، روند از داده‌ها باید حذف شود. حذف شدن روند باعث می شود میانگین داده‌ها ثابت بماند. لگاریتم‌گیری از متغیر ها یکی دیگر از روشهایی است که به ایستایی کمک نموده و باعث تغییر مقیاس در داده‌ها می‌شود [۲]. ضمن اینکه لگاریتم‌گیری از متغیرها هم‌خطی موجود بین متغیرها را نیز از بین برده و باعث افزایش در دقت مدل سازی می‌شود. در نتیجه در حالت سوم ابتدا از داده‌ها لگاریتم گرفته شده و سپس روند موجود از آنها حذف می شود و در نهایت داده‌ها نرمال می‌شوند. بعد از پیش‌بینی مقادیر تولید پسماند برای آینده، ترم‌هایی که از آنها حذف شده به آنها اضافه می‌شود. تفاضل‌گیری نیز یکی از روش‌های مانا کردن سری زمانی است. چون بعد از پیش‌بینی نمی‌توان مقادیری که خروجی شبکه عصبی است به واحد تولید پسماند تبدیل نمود از این روش استفاده نشده‌است (در این حالت مقادیر خروجی از جنس تفاضل گیری شده می باشد).

۷- جمع بندی

حالت اول: در این قسمت کلیه داده های متغیرهای وابسته و مستقل دوره مشاهداتی و متغیرهای مستقل دوره‌ی پیش بینی بدون اعمال هرگونه تغییری در برنامه شبکه عصبی وارد شده اند.

با توجه به معیارهای انتخاب مدل بهترین شبکه عصبی در این قسمت مشخص شد. در شبکه عصبی انتخاب شده از traingdm, learngdm استفاده شده‌است و معماری شبکه بدین صورت است که دارای ۲ لایه میانی باشد که در لایه اول ۵ نرون و در لایه دوم ۶ نرون موجود می باشد. برای ارزیابی دقت مدل از سه پارامتر عملکرد بهینه MSE، MAPE و ضریب همبستگی استفاده شده است. مقادیر این آماره‌ها به ترتیب ۰/۴۶ ،۰/۰۶ و ۰/۷۴ بدست آمده است.

حالت دوم: در این قسمت از کلیه داده‌های متغیر وابسته و مستقل دوره‌ی مشاهداتی و متغیرهای مستقل دوره‌ی پیش‌بینی روند حذف شده‌است. روند موجود در مقادیر مشاهداتی پسماند شهری برابر است با:

$$Y = 104/08\,t + 39375 \qquad (۲)$$

که در آن:

T: زمان بر حسب ماه

Y: پسماند تولیدی

نتایج آزمون دیکی فولر نشان داد، بعد از حذف روند، نتایج آزمون دیکی فولر بهتر شده و سری‌های زمانی ایستا تر شده اند.

با توجه به معیارهای انتخاب مدل بهترین شبکه عصبی در این قسمت مشخص شد. در شبکه عصبی انتخاب شده از traincgb, learngdm استفاده شده است و معماری شبکه بدین صورت است که دارای ۲ لایه میانی که در لایه اول ۹ نرون و در لایه دوم ۱۰ نرون موجود می‌باشد.

برای ارزیابی دقت مدل از سه پارامتر عملکرد بهینه MAPE، MSE و ضریب همبستگی استفاده شده‌است. مقادیر این آماره‌ها به ترتیب ۰/۰۵، ۰/۴۳ و ۰/۷۹ بدست آمد.

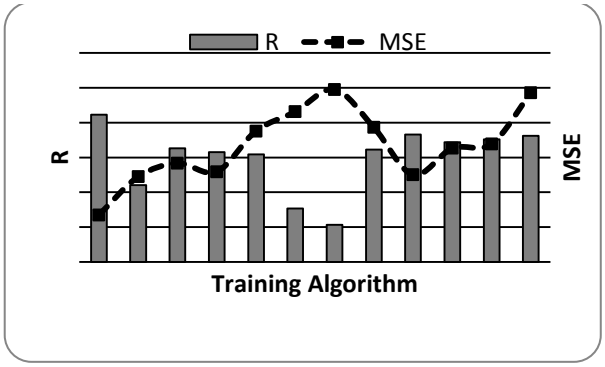
مقادیر این آماره‌ها نشان دهنده دقت خوب در مدل سازی می-باشد.

حالت سوم: برای اینکه مقیاس داده‌های ورودی در مرحله پیش‌بینی به مقادیر استفاده شده در مرحله آموزش نزدیکتر شود ابتدا از داده‌ها لگاریتم گرفته می‌شود و روند موجود در متغیرها حذف می‌شوند. لگاریتم‌گیری ضمن اینکه باعث می‌شود سری داده‌ها ملایم‌تر شوند، به ایستا شدن یک سری زمانی کمک کرده و دقت مدل سازی را بالا می‌برد. نتایج آزمون دیکی فولر نشان داد

که لگاریتم گیری توانسته است مقادیر این آزمون را بهبود بخشد.

با توجه به معیارهای انتخاب مدل، بهترین شبکه عصبی در این قسمت مشخص شد. در شبکه عصبی انتخاب شده از train gda, learngdm استفاده شده است و دارای ۲ لایه میانی بوده که در لایه اول ۵ و لایه دوم ۵ نرون موجود می‌باشد. برای ارزیابی دقت مدل از سه آزمون MAPE، MSE و ضریب همبستگی استفاده شده‌است. مقادیر این آماره‌ها به ترتیب ۰/۰۴۶، ۰/۲۶ و ۰/۸۶ بدست آمده است.

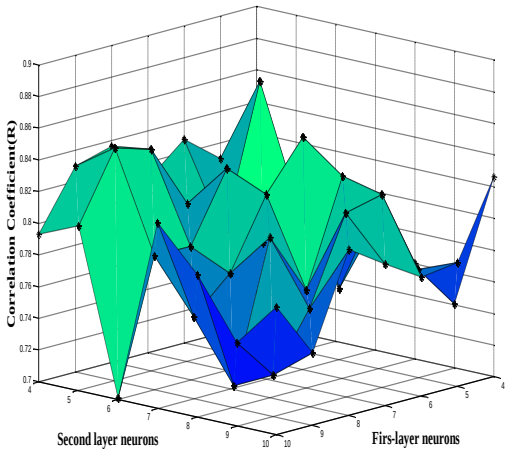
شکل(۲) مقایسه‌ای بین الگوریتم‌های گوناگون را برای شبکه BPNN چهار لایه با ۳ نورون در لایه ورودی، ۵ نورون در لایه پنهان اول و دوم و یک نورون در لایه پنهان خروجی (ساختار ۳:۵:۵:۱) نشان می‌دهد. همان‌طور که دیده می شود، بیش‌ترین ضریب تعیین و کمترین خطا از آن الگوریتم (traingda) است.



شکل (۲): تأثیر الگوریتم آموزش روی میزان خطا و ضریب تعیین شبکه BPNN

شکل (۳) و (۴) نمودار سه بعدی میزان ضریب همبستگی، تعداد نرون‌های لایه اول و تعداد نرون‌های لایه دوم به طور واضح نشان می‌دهند که شبکه عصبی در حالت سوم در حالت ۳:۵:۵:۱ بیشترین R و کمترین خطا را دارد.

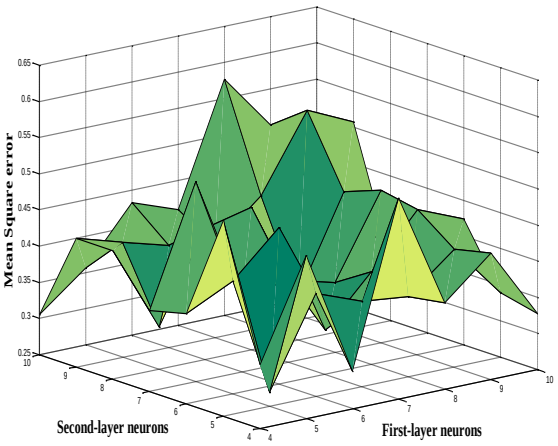
در این تحقیق اثر پیش پردازش متغیرهای سری زمانی بر روی مدلسازی با شبکه عصبی بررسی شد.



شکل (۳): نمودار سه بعدی میزان خطا برحسب تعداد نرون در لایه اول و لایه دوم

۸- نتیجه‌گیری

از آنجا که شبکه‌های عصبی درونیاب بوده و در محدوده اطلاعات ورودی، آموزش می‌بینند قادر نیستند مقادیر تغییرات زیاد متغیرها را در دراز مدت ببینند و بنابراین متغیر خروجی از شبکه در مرحله پیش‌بینی در محدوده خروجی در مرحله آموزش بدست می‌آید. به منظور حل این مشکل سعی شد داده‌ها به حالت ایستایی نزدیک شوند و مقیاس داده‌های ورودی و خروجی در مرحله آموزش و مقیاس داده‌های ورودی در مرحله پیش‌بینی کوچک شود. بدین منظور با حذف روند و پیش بینی بر اساس جز تصادفی، متغیرهای سری زمانی به حالت ایستایی نزدیک‌تر شده و مقیاس داده‌ها کوچکتر می‌شوند. علاوه بر این لگاریتم گیری باعث افزایش دقت مدلسازی شده و سری داده‌ها را به حالت ایستایی نزدیک می نماید. نتایج نشان داد با لگاریتم گیری از متغیرهای سری زمانی ورودی و خروجی و حذف روند داده‌ها، رسیدن به حالت ایستایی محتمل‌تر می‌شود و مدل‌سازی دقیق‌تر انجام می-گیرد.



شکل(۴): نمودار سه بعدی میزان همبستگی برحسب تعداد نرون در لایه اول و لایه دوم

۹- مراجع

[۱] عبدلی، م.ع.، "مدیریت مواد زائد جامد شهری" ، انتشارات سازمان شهرداری های کشور، جلد اول، ۱۳۷۹.

[۲] Beigl,P.; Wassermann,G.; Schneider,F., and Salhofer,S. "Forecasting municipal solid waste generation in major European cities", In:Pahl Wostl, C., Schmidt, S.,Jakeman, T.(Eds.), iEMSs 2004 International Congress: Complexity and Integrated Resources Management. Osnabrueck, Germany, 2004.

[۳] Bach, H.; Mild, A.; Natter, M.; Weber, A. "Combining socio-demographic and logistic factors to explain the generation and collection of waste paper", Resources Conservation and Recycling, vol.41, pp. 65– 73, 2004.

[۴] Chung, s. "Projecting municipal solid waste: The case of domestic waste in HongKong special administration region", Environmental enginerring science, vol. 27, pp. 13- 20, 2010.

[۵] Chung.s. "Projection of trends in solid waste generation:The case of HongKong SAR". Resource, conservation and recycling. Vol. 54, pp.759- 768, 2010.

[۶] Daskalopoulos, E.; Badr, O., and Probert. S.D. "Municipal solidwaste: A prediction methodology for the generation rate and composition in the European Union countries and the United States of America", Resources, Conservationand Recycling, vol.24, pp.155– 166, 1998.

[۷] Dyson, B.; Chang, N. "Forecasting municipal solidwaste generation in a fast-growin Urban region with system dynamics modeling", WasteManagement, vol. 25, pp. 669– 679, 2005.

[۸] Iffat, A.; Leslie, S. "A Neural Network Approach to Time Series Forecasting", proceeding of the world congress on engineering,vol. 2, july, pp. 1-3, London,uk, 2009.

[۹] Ojeda Benítez, S. "Mathematical modeling to predict residential solid waste generation", Waste Management, vol. 28, pp. S7– S13, 2008.

[۱۰] Noori,R.;Abdoli, M.; Jalili Ghazizade, M.; Samieifard,R. "Comparison of Neural Network and Principal Component-Regression Analysis to Predict the Solid Waste Generation in Tehran", Iranian J Publ Health, Vol.38, pp. 74- 84, 2009.

[۱۱] Noori, R., Abdoli, M.A.; AmeriGhasrodashti, A., JaliliGhazizade, M. "Prediction of Municipal Solid Waste Generation with Combination of Support Vector Machine and Principal Component Analysis: A Case Study of Mashhad", Environmental Progress &Sustainable Energy, vol.28, pp. 249- 258, 2008.

[۱۲] okka,L., Antikainen,R., Kauppi, P; "Municipal solid waste production and composition in Finland Changes in the period 1960– 2002 and prospects until 2020", Resources, Conservation and Recycling, vol. 50, pp. 475– 488, 2007.

[۱۳] Tawfiq, A.; Ibrahim El, Amin., "Artificial neural networks as applied to long-term demand forecasting", Artificial Intelligence in Engineering, vol.13, pp. 189– 197, 2009.

[۱۴] Zhang, G. "Timeseries forecasting using a hybrid ARIMA and neural network model", Neurocomputing, vol. 50, pp. 159– 175, 2003.

۱۰- زیرنویس ها

[۱] Daskalopoulos
[۲] Beigl
[۳] Dyson
[٤] Stella
[٥] Sokka
[٦] Sara Ojeda Benítez
[۷] Bach
[۸] Chung
[۹] MATHLAB