

## A New Developed Model for Cascaded Optimal Neural Networks for Short Term Electricity Demand Forecasting

Hossein Jalali <sup>1\*</sup>, Faezeh Amirteimoury<sup>2</sup>

<sup>1</sup>Technical Expert at Ramin Power Generation Company, Khoozestan, Ahvaz. Email: [hossein.jalali11@gmail.com](mailto:hossein.jalali11@gmail.com)

<sup>2</sup>South of Kerman Electricity Distribution Company, Kerman. Email: [famirteimoury@gmail.com](mailto:famirteimoury@gmail.com)

### ARTICLE INFO

**Article type:**  
Research Paper

**Article History:**  
Received 19 May 2025  
Revised 23 June 2025  
Accepted 13 September 2025  
Published Online 23 September 2025

**Keywords:**  
Short-term load forecasting,  
feature selection,  
feed-forward neural network,  
composite neural network,  
particle swarm optimization,  
power systems.

### ABSTRACT

Short-term load forecasting (STLF) is precisely required for planning, operation, and control of the power system. This forecasting method is used by electricity installations, system operators, generators, and energy marketers. In this paper, by applying the feature selection (FS) technique on the basis of the Mutual Information (MI) criterion, the most effective of these data are selected. The proposed method in this paper is to use the Composite neural network (CNN), which consisting of three feedforward neural networks of the type multi-layer perceptron that are connected in series. Also, the Particle swarm optimization method (PSO) is used to achieve optimal values of the impact parameters in the forecasting process, which are included in the feature selection section and neural network. this method has been applied to load data of year 97 cities of Kerman and its results emphasize the efficiency of this method in the forecasting of electric load.

**Cite this article:** Jalali, H., & Amirteimoury, F. (2025). A New Developed Model for Cascaded Optimal Neural Networks for Short Term Electricity Demand Forecasting. *Energy Systems: Technology and Management*, 1 (1), 58-69. DOI: <http://doi.org/10.48306/estm.2025.232464>



© Graduate University of Advanced Technology.  
DOI: <http://doi.org/10.48306/estm.2025.232464>

\* Corresponding author.  
Email: [mohammad\\_kakuei@yahoo.com](mailto:mohammad_kakuei@yahoo.com)



## پیش بینی تقاضای کوتاه مدت انرژی الکتریکی بر اساس یک مدل بهینه توسعه یافته آموزش شبکه های عصبی ترکیبی

حسین جلالی<sup>۱\*</sup>، فائزه امیرتیموری<sup>۲</sup>

<sup>۱</sup> کارشناس فنی در شرکت تولید برق رامین، خوزستان، اهواز. رایانامه: [hossein.jalali11@gmail.com](mailto:hossein.jalali11@gmail.com)

<sup>۲</sup> شرکت توزیع برق جنوب کرمان، کرمان. رایانامه: [famirteimoury@gmail.com](mailto:famirteimoury@gmail.com)

چکیده	اطلاعات مقاله
پیش بینی کوتاه مدت بار (STLF) به طور دقیق، برای برنامه ریزی، بهره برداری و کنترل سیستم قدرت مورد نیاز است. این روش پیش بینی توسط تاسیسات برق، اپراتورهای سیستم، ژنراتورها و بازارهای انرژی استفاده می شود. در این مقاله با اعمال تکنیک انتخاب ویژگی (FS) بر مبنای معیار اطلاعات متقابل (MI) بر روی داده های بار الکتریکی ورودی، موثرترین این داده ها انتخاب می شوند. روش پیشنهادی در این مقاله استفاده از شبکه عصبی ترکیبی (CNN) که متشکل از سه شبکه عصبی پیش خور از نوع پرسپترون چند لایه (MLP) که به صورت سری به یکدیگر متصل شده اند، می باشد. همچنین از روش بهینه سازی ازدحام ذرات (PSO) به منظور دستیابی به مقادیر بهینه پارامترهای تاثیر گذار در روند پیش بینی، که این پارامترها در قسمت انتخاب ویژگی و شبکه عصبی پیشنهادی قرار دارند، استفاده می شود. این روش پیشنهادی بر روی داده های بار سال ۹۷ شهرستان کرمان اعمال شده و نتایج آن تأییدی بر کارایی این روش در پیش بینی بار الکتریکی می باشد.	<b>نوع مقاله:</b> پژوهشی <b>تاریخ های مقاله:</b> تاریخ دریافت: ۱۴۰۴/۰۲/۲۹ تاریخ بازنگری: ۱۴۰۴/۰۴/۰۲ تاریخ پذیرش: ۱۴۰۴/۰۶/۲۲ تاریخ انتشار: ۱۴۰۴/۰۷/۰۱ <b>کلیدواژه:</b> پیش بینی بار کوتاه مدت، انتخاب ویژگی، شبکه عصبی پیش خور، شبکه عصبی ترکیبی، بهینه ساز ازدحام ذرات.
جلالی، حسین و امیرتیموری، فائزه (۱۴۰۴). پیش بینی تقاضای کوتاه مدت انرژی الکتریکی بر اساس یک مدل بهینه توسعه یافته آموزش شبکه های عصبی ترکیبی. سیستم‌های انرژی: فناوری و مدیریت، ۱ (۱) ۶۹-۵۸. DOI: <a href="http://doi.org/10.48306/estm.2025.232464">http://doi.org/10.48306/estm.2025.232464</a>	<b>استناد:</b>



## ۱- مقدمه

از شبکه های عصبی با نام شبکه عصبی ترکیبی (CNN) در این مقاله معرفی شده است، که نسبت به سایر روش های گفته شده دارای قدرت بیشتر در پیش بینی پارامترهای غیر خطی سیستم های قدرت می باشد و خطای تخمین لحظات آینده پارامترهای سیستم قدرت را به شدت کاهش می دهد و دقیق ترین تخمین را ارائه می دهد.

همچنین از روش بهینه سازی با الگوریتم ازدحام ذرات (PSO) برای محاسبه پارامترهای استفاده شده در مرحله انتخاب ویژگی و مرحله شبکه عصبی پیشنهادی استفاده شده است. مدل پیشنهادی مورد مطالعه در این مقاله از ترکیب شبکه عصبی ترکیبی و روش بهینه سازی ازدحام ذرات حاصل می شود که نتایج حاکی از قدرت و کارایی این مدل در پیش بینی بار الکتریکی می باشد.

این مدل پیشنهادی بر روی داده های بار الکتریکی برای سال ۹۷ شهرستان کرمان اعمال شده که این داده ها به صورت ساعتی در طول یک سال ثبت شده اند که مجموعاً ۸۷۶۰ داده از نوع بار می باشند.

## ۲- نرمال کردن داده های بار

داده های بار اعداد نسبتاً بزرگی هستند و تجزیه و تحلیل این داده ها بسیار سخت و زمان زیادی را می طلبد، برای همین با استفاده از رابطه ۱، این داده ها در بازه صفر و یک محدود شده اند تا زمان تجزیه و تحلیل را کمتر کنند و کیفیت عملکرد تجزیه و تحلیل داده ها را افزایش دهند تا شبکه به راحتی بتواند الگوها را یاد بگیرد و نتایج خروجی بهتری را ایجاد کند.

$$L_i = \frac{L_i - \min(L)}{\max(L) - \min(L)} \quad (1)$$

$L_i$  داده بار  $i$  ام و  $i = 1, 2, 3, \dots, 8760$  می باشد.

## ۳- آماده سازی داده های بار الکتریکی

منظور از آماده سازی داده های بار الکتریکی، در واقع ساختن ماتریس داده های ورودی اولیه و بردار هدف جهت ورود به شبکه عصبی و آموزش دادن آن ها و انجام عمل پیش بینی بر روی این داده ها می باشد که به ترتیب در ۲ و ۳ مشاهده می شوند. دقت شود که بردار هدف در واقع همان بار ساعات بعدی و هر سطر از ماتریس داده های ورودی اولیه یک متغیر ورودی (IV) است که شماره متغیر ورودی معرف داده های ساعت قبل از بردار هدف می باشد.

به عنوان مثال،  $IV_1$  نمایانگر داده های مربوط به یک ساعت قبل از مقدار بردار هدف است و  $IV_2$  نیز به داده های دو ساعت قبل اشاره دارد. بنابراین، هر سطر از ماتریس داده های ورودی اولیه میزان شیفت به ساعات قبل (Back Shift) را نشان می دهد که در این مقاله مقدار آن برابر با ۲۰۰ (BS=200) در نظر گرفته شده است. همچنین، ستون های این ماتریس و بردار هدف معرف نمونه های آموزشی هستند که هر یک، داده های یک ساعت قبل از مقدار هدف را شامل می شوند. در این تحقیق، تعداد نمونه های آموزشی برابر با ۱۵۰۰ (TS=1500) در نظر گرفته شده است.

پیش بینی بار، به خصوص پیش بینی بار کوتاه مدت (STLF) نقش بسیار مهمی در زمینه مدیریت سیستم های انرژی و برنامه ریزی بهتر این سیستم ها ایفا می کند. این نوع پیش بینی، عملکرد قابل اعتماد سیستم قدرت را تضمین می کند که منجر به تامین بی وقفه برق برای مصرف کننده می شود [۱]. همچنین برنامه ریزی، نگهداری، تنظیم نرخ تعرفه ها و ارزیابی قرارداد را می توان به راحتی با پیش بینی بار دقیق انجام داد [۲، ۳]. یکی دیگر از کاربردهای مهم پیش بینی بار کوتاه مدت، توزیع اقتصادی و قابلیت اطمینان سیستم قدرت است. قابلیت اطمینان سیستم قدرت با تغییرات ناگهانی تقاضای بار در نوسان است. به عنوان مثال، اگر تقاضای بار کم شود، ممکن است سیستم با کمبود عرضه برق مواجه شود، بنابراین مدیریت شرایط اضافه بار و کیفیت کلی سیستم قدرت دشوار است. بالعکس، اگر تقاضای بار بیش از حد مورد نیاز تخمین زده شود، مقدار زیادی از منابع موجود بیشتر صرف خواهند شد. در چنین شرایطی، دقت بار در دهه گذشته محبوبیت فوق العاده ای به دست آورده است [۴].

به طور کلی روش های متعددی برای پیش بینی بار وجود دارد، اما عمده ترین تکنیک ها را می توان به دو دسته تقسیم کرد:

۱. تکنیک های مبتنی بر روش های آماری

۲. تکنیک های مبتنی بر هوش مصنوعی (AI) [۵، ۶].

روش های آماری شامل تکنیک های سری زمانی همچون رگرسیون خطی چندگانه [۷]، رگرسیون پویا و تابع انتقال [۸، ۹]، اتورگرسیون میانگین متحرک [۱۰]، ناهمسانی واریانس شرطی اتورگرسیون تعمیم یافته [۱۱] می باشند. همچنین تکنیک های هموارسازی نمایی [۱۲] و فیلترهای کالمن [۱۳] از دیگر تکنیک های مبتنی بر روش های آماری می باشند. در روش های آماری خطای پیش بینی کمی خواهیم داشت اگر رفتار داده های ورودی در شرایط عادی باشند، یعنی اگر در داده های ورودی تغییرات ناگهانی وجود داشته باشد، نمی توان این تغییرات را ردیابی کرد و خطای پیش بینی خیلی بالایی در خروجی این روش خواهیم داشت.

تکنیک های پیش بینی مبتنی بر هوش مصنوعی مانند شبکه عصبی مصنوعی، منطق فازی و بردار پشتیبان ماشین [۱۹-۱۴] در مسائل پیش بینی مورد استفاده قرار می گیرند که شبکه عصبی مصنوعی عملکرد بهتری را در مقایسه با دیگر تکنیک های ذکر شده برای متغیرهای ورودی غیر خطی ارائه می دهد، همچنین دارای توانایی بهتری در ایجاد نگاهشست بین داده های ورودی و خروجی بدون ایجاد فرمول بندی پیچیده ریاضی دارد.

شبکه عصبی مصنوعی دارای ساختار متنوعی می باشد که هر کدام از این ساختارها در زمینه های مختلف مورد استفاده قرار می گیرند. شبکه عصبی پیشنهادی در این مقاله، حاصل ترکیب متوالی سه شبکه عصبی یکسان دارای ساختار شبکه عصبی پیش خور (FNN) [۲۰] از نوع پرسپترون چند لایه (MLP) می باشد که یک مدل توسعه یافته

با کمک معیار اطلاعات متقابل به محاسبه همبستگی بین متغیرهای ورودی پرداخته می‌شود.

#### ۴-۱- فیلتر داده‌های کم اثر

فرض کنید  $S_{total} = \{IV_1, IV_2, IV_3, \dots, IV_{2000}\}$  مجموعه‌ای از ویژگی‌ها یا همان متغیرهای ورودی که نامزدهای ورودی به قسمت انتخاب ویژگی هستند و  $\gamma$  نیز بردار هدف باشد. در این صورت برای هر ویژگی  $IV_i \in S$  که  $i = 1, 2, 3, \dots, 2000$ ، معیار MI آن با بردار هدف ( $\gamma$ ) به صورت  $MI(IV_i, \gamma)$  نشان داده می‌شود و طریقه محاسبه آن بر اساس محاسبه توزیع احتمالی مشترک  $P(IV_i, \gamma)$  و توزیع احتمالی مستقل  $P(IV_i)$  و  $P(\gamma)$  می‌باشد. توجه شود که هر چقدر معیار  $MI(IV_i, \gamma)$  بزرگ باشد نشان دهنده همبستگی بالای این دو ویژگی می‌باشد. برای تعیین مقدار بهینه میزان همبستگی بین این دو ویژگی از پارامتر  $TH_1$  که آستانه فیلتر داده‌های کم اثر نام دارد، استفاده می‌شود.

$$MI(IV_i, \gamma) \geq TH_1 \quad (5)$$

معادله (۵) بیان می‌کند که فقط ویژگی‌هایی که دارای همبستگی بزرگتر یا مساوی  $TH_1$  هستند حق عبور از این فیلتر را دارند و بقیه آن‌ها به دلیل کم اثر بودن حذف می‌شوند. نحوه بدست آوردن مقدار بهینه  $TH_1$  در این مقاله از طریق روش بهینه سازی ازدحام ذرات می‌باشد که در قسمت‌های بعدی توضیح داده می‌شود.

#### ۴-۲- فیلتر افزونگی یا داده‌های تکراری

فرض کنید که  $S_1 \subset S_{total}$  باشد، که زیر مجموعه‌ای از ویژگی‌ها یا همان متغیرهای ورودی که حاصل از فیلتر کردن داده‌های نامربوط توسط فیلتر داده‌های کم اثر می‌باشند. در این مرحله نیز برای هر  $IV_{i,j} \in S_1$  معیار MI آن برای متغیر ورودی به نام  $i$  متغیر ورودی به نام  $j$  که  $i \neq j$  است به صورت  $MI(IV_i, IV_j)$  نشان داده می‌شود و طریقه محاسبه آن بر اساس محاسبه توزیع احتمال مشترک  $P(IV_i, IV_j)$  و توزیع احتمال مستقل  $P(IV_i)$  و  $P(IV_j)$  می‌باشد. در واقع در این قسمت همبستگی هر سطر از ماتریس داده‌های عبوری از مرحله اول با سطر دیگر بررسی می‌شود و بر اساس معیار MI تصمیم‌گیری صورت می‌گیرد. هر چقدر که معیار  $MI(IV_i, IV_j)$  بزرگ باشد نشان دهنده همبستگی زیاد و تشابه ویژگی‌ها نسبت به یکدیگر می‌باشد که در این مرحله نیاز است که یکی از این ویژگی‌ها یا همان متغیر ورودی حذف شود، زیرا به دو ویژگی شبیه به همدیگر احتیاج نداریم و با حذف ویژگی‌های مشابه، علاوه بر کاهش ابعاد داده‌ها، باعث عملکرد بهتر مرحله پیش‌بینی و بدست آوردن نتایج بهتر و دقیق‌تر می‌شود. برای تعیین مقدار بهینه میزان همبستگی بین ویژگی‌ها از پارامتر  $TH_2$  که آستانه فیلتر داده‌های تکراری نام دارد، استفاده می‌شود.

$$MI(IV_i, IV_j) \geq TH_2 \quad (6)$$

در این مقاله پیش‌بینی بار الکتریکی برای هر فصل از سال به صورت جداگانه صورت گرفته است در نتیجه برای هر فصل باید ماتریس داده‌های ورودی اولیه و بردار هدف تشکیل شود. لازم است که تعداد نمونه‌های انتخاب شده برای هر فصل از سال، بزرگتر یا مساوی مجموع نمونه‌های آموزشی (TS) و میزان شیفت به ساعات قبل (BS) در نظر گرفته شود.

برای انجام عمل پیش‌بینی بر روی داده‌ها، لازم است که داده‌های ماتریس داده‌های ورودی اولیه و بردار هدف را به دو دسته داده‌های آموزشی و داده‌های آزمایشی تقسیم‌بندی کرد. در این مقاله ۵۰۰ داده ابتدایی را به عنوان داده‌های آزمایشی و بقیه داده‌ها به عنوان داده‌های آموزشی در نظر گرفته شدند.

$$Target_{1-1500}(\gamma) = [x_{h-1499}, \dots, x_{h-2}, x_{h-1}, x_h] \quad (2)$$

$$InputMatrix_{200 \times 1500} = \begin{pmatrix} IV_1 & x_{h-1501}, \dots, x_{h-3}, x_{h-2}, x_{h-1} \\ IV_2 & x_{h-1002}, \dots, x_{h-4}, x_{h-3}, x_{h-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ IV_{200} & x_{h-1699}, \dots, x_{h-202}, x_{h-201}, x_{h-200} \end{pmatrix} \quad (3)$$

$$InputMatrix(200 \times 1500) = \begin{matrix} IV_1 & x_{h-1500} & \dots & x_{h-2} & x_{h-1} \\ IV_2 & x_{h-(1501)} & \dots & x_{h-3} & x_{h-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ IV_{200} & x_{h-(1699)} & \dots & x_{h-(201)} & x_{h-200} \end{matrix} \quad (4)$$

#### ۴-۳ انتخاب ویژگی

همان‌طور که در ماتریس داده‌های ورودی اولیه (۳) دیده می‌شود، این ماتریس دارای داده‌های بسیار زیاد می‌باشد که وارد کردن این داده‌ها به طور مستقیم به قسمت پیش‌بینی یک عمل کاملاً غیر اصولی می‌باشد، زیرا در بین این داده‌ها قطعاً یک سری داده‌های غیر موثر و تکراری وجود دارند که وارد کردن آن‌ها به قسمت پیش‌بینی باعث کاهش دقت و افزایش خطای پیش‌بینی می‌شود.

تمرکز مسئله انتخاب ویژگی (FS) بر روی یافتن یک زیرمجموعه از ورودی‌های ارزنده می‌باشد. در حقیقت این عمل منجر به بهبود عملکرد یادگیری، کاهش پیچیدگی محاسباتی، کاهش فضای ذخیره‌سازی و ساخت مدلی بهتر با قابلیت تعمیم می‌شود [۲۱]. در این بخش به معرفی روشی برای فیلتر کردن داده‌ها در دو مرحله بر مبنای معیار اطلاعات متقابل (MI) که این روش مبتنی بر تابع آنترپی می‌باشد و توضیحات کافی در مورد این روش در [۲۲، ۲۳] داده شده است.

منظور از فیلتر کردن داده‌ها در دو مرحله:

۱. فیلتر داده‌های کم اثر:

با کمک معیار اطلاعات متقابل به محاسبه میزان همبستگی بین مؤلفه‌های متغیر ورودی (IV) و مؤلفه‌های بردار هدف پرداخته می‌شود.

۲. فیلتر داده‌های تکراری:

شبکه‌های عصبی مورد استفاده در شبکه عصبی پیشنهادی، دارای یک خروجی هستند که در واقع همان مقدار بار پیش‌بینی شده برای ساعت آینده می‌باشد. از هر شبکه عصبی موجود در شبکه عصبی ترکیبی، وزن‌ها و بایاس‌های آن شبکه به شبکه‌ی بعدی منتقل می‌شوند، به عبارت دیگر هر شبکه عصبی، دانش کسب شده و ذخیره شده در وزن‌ها و بایاس‌های خود را به شبکه عصبی بعدی منتقل می‌کند. بنابراین می‌توان فرآیند آموزش خود را از نقطه پایان شبکه عصبی قبلی آغاز کند. فقط اولین شبکه عصبی موجود در شبکه عصبی ترکیبی باید با مجموعه مقادیر اولیه وزن‌ها و بایاس‌های تصادفی شروع شود و بقیه شبکه‌های عصبی، وزن‌ها و بایاس‌های آموزش دیده‌ی شبکه عصبی قبل از خود را دریافت می‌کنند، که این کار موجب کاهش محدودیت در آموزش و عملکرد بهتر شبکه عصبی در روند آموزش می‌شود. هر شبکه عصبی موجود در شبکه عصبی ترکیبی، دارای الگوریتم آموزشی مختص به خود آن شبکه عصبی می‌باشد که فرآیند آموزش بر روی داده‌ها را انجام می‌دهد. مدل پیشنهادی این مقاله شامل سه لایه مجزا بوده که لایه اول با استفاده از الگوریتم LM آموزش داده شده و به بهینه سازی گردید. سپس وزن و بایاس های تولید شده در لایه اول به لایه دوم منتقل شده و با استفاده از الگوریتم BFGS آموزش داده شد. در نهایت مقادیر وزن و بایاس لایه دوم به لایه سوم که با استفاده از الگوریتم LM آموزش داده می‌شود، منتقل شده و خروجی نهایی که یک مدل پیش بینی کننده می‌باشد، بدست می‌آید. ترکیب استفاده شده یعنی استفاده از الگوریتم LM به عنوان مدل آموزش در لایه اول و آخر و همچنین استفاده از الگوریتم BFGS به عنوان مدل آموزش لایه سوم منجر به بهبود عملکرد کلی شبکه عصبی ترکیبی و کسب نتایج بهتر می‌شود [۲۴].

یادگیری شبکه عصبی، براساس خطای ایجاد شده در طول فرآیند آموزش است. بنابراین، خطای شبکه به صورت مربع میانگین اختلاف بین مقادیر خروجی هدف و خروجی شبکه عصبی تعریف می‌شود. این خطای میانگین مربع (MSE) شبکه عصبی است که در طول فرآیند آموزش شبکه مورد استفاده قرار می‌گیرد. به طور معمول، خطای میانگین مربع به عنوان خطای یادگیری شبکه و مقادیر وزنی شبکه عصبی برای رسیدن به هدف یادگیری مورد نظر، برورسانی می‌شوند. خطای آموزش دیگری نیز مورد استفاده قرار می‌گیرد که به آن جذر خطای میانگین مربع (RMSE) گفته می‌شود. برای سنجش خطای تست از خطای درصدی متوسط قدرمطلق (MAPE) استفاده می‌شود.

$$MSE = \frac{1}{N} \sum_{i=1}^N (F(i) - Target(i))^2 \quad (7)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (F(i) - Target(i))^2} \quad (8)$$

$$MAPE = 100 * \frac{1}{N} \sum_{i=1}^N \frac{|(F(i) - Target(i))|}{Target(i)} \quad (9)$$

معادله (۶) بیان می‌کند که با بزرگ‌تر یا مساوی بودن معیار  $MI(IV_i, IV_j)$  از  $TH_2$  یکی از ویژگی‌ها باید حذف شود و فقط یکی از آن دو ویژگی حق عبور از این فیلتر را دارد. نحوه بدست آوردن مقدار بهینه  $TH_2$  در این مقاله از طریق روش بهینه سازی ازدحام ذرات می‌باشد که در قسمت‌های بعدی توضیح داده می‌شود. در نهایت  $S_2 \subset S_1$  که زیرمجموعه ویژگی‌های نهایی می‌باشد، به عنوان خروجی از الگوریتم انتخاب ویژگی دو مرحله‌ای و ورودی به شبکه عصبی ترکیبی انتخاب شده و ویژگی‌های موجود در  $S_2$  را با نام ماتریس داده‌های ثانویه در این مقاله مورد استفاده قرار می‌گیرند.

به این نکته باید توجه کرد که هر چه مقدار  $TH_1$  را افزایش یابد، داده‌های بیشتری حذف می‌شوند و باند عبور باریکتر می‌شود و فقط داده‌هایی که بیشترین اثر را دارند و همبستگی آن‌ها نسبت به هدف، در بالاترین حد باشند از این باند عبور می‌کنند.

همچنین هر چه مقدار  $TH_2$  کمتر شود، داده‌های بیشتری حذف می‌شوند و باند عبور باریکتر می‌شود و فقط داده‌هایی که کمترین شباهت را با هم دارند از این باند عبور می‌کنند، در واقع داده‌هایی که شبیه هم هستند را به عنوان داده‌های اضافی و تکراری، آن‌ها را فیلتر می‌کند.

## ۵- شبکه عصبی ترکیبی (Cascade Neural Network)

همان طور که قبلاً گفته شد، بار الکتریکی یک تابع غیر خطی و متغیر با زمان می‌باشد و برای یک شبکه عصبی، ایجاد نگاشت صحیح بین ورودی و خروجی از چنین سیگنالی در تمام دوره‌های زمانی بسیار مشکل است و این امر به دلیل وجود محدودیت در آموزش برای یک تک شبکه عصبی می‌باشد. به همین علت در این مقاله از شبکه عصبی ترکیبی، متشکل از سه شبکه عصبی متوالی که هر شبکه دارای ساختار شبکه عصبی پیش خور از نوع پرسپترون چند لایه (MLP) که هر شبکه دارای الگوریتم یادگیری به خصوصی می‌باشد، استفاده شده و این شبکه عصبی پیشنهادی به خوبی از عهده محدودیت موجود در آموزش بر می‌آید. در ساختار شبکه MLP، تعداد نورون‌های لایه ورودی و لایه پنهان و لایه خروجی برای هر شبکه از شبکه عصبی پیشنهادی، یکسان در نظر گرفته شده، دقت شود که در ساختار MLP می‌توان از چند لایه پنهان استفاده کرد ولی در این مقاله از یک لایه پنهان در هر شبکه استفاده شده است. تعداد نورون‌های لایه پنهان (NH) در هر شبکه بر خلاف تعداد نورون‌های لایه ورودی و لایه خروجی که مقدار ثابت و مشخصی دارند، می‌توان هر مقداری به آن اختصاص داد. در واقع نورون‌های لایه پنهان (NH)، از نوع پارامتر قابل تنظیم می‌باشد که برای انتخاب بهینه‌ترین مقدار برای آن، از روش بهینه سازی ازدحام ذرات استفاده می‌شود که در قسمت‌های بعدی توضیح داده شده است. نامزدهای ورودی به شبکه عصبی ترکیبی همان مؤلفه‌های انتخاب شده در قسمت انتخاب ویژگی که در واقع همان ماتریس داده‌های ورودی ثانویه می‌باشند. دقت شود که تمامی

$$\omega^t = \omega_{max} - \frac{(\omega_{max} - \omega_{min})}{t_{max}} \times t \quad (11)$$

که مقدار  $\omega_{min} = 0.4$  و  $\omega_{max} = 0.9$  در نظر گرفته شده است.

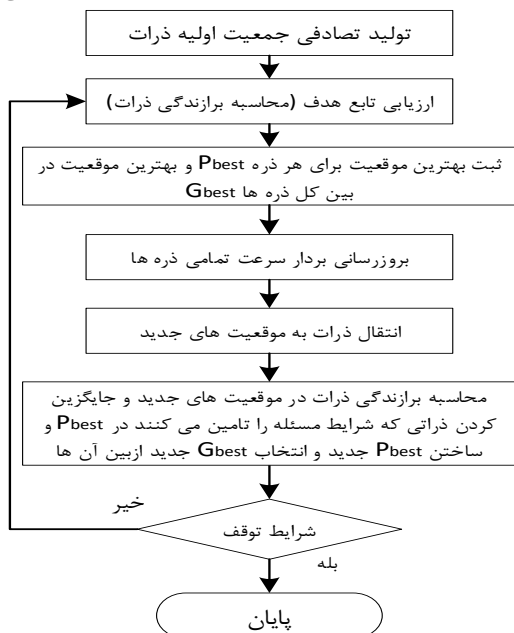
۵. موقعیت هر ذره به صورت زیر بروزرسانی می‌شود:

$$x_i^{t+1} = v_i^{t+1} + x_i^t \quad (12)$$

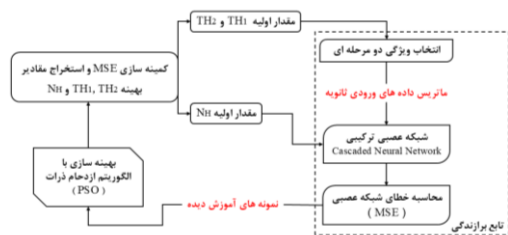
۶. در مرحله ۴ و ۵ به ترتیب  $v$  و  $x$  بروزرسانی می‌شوند، همچنین تابع برازندگی مربوط به  $x$  جدید محاسبه می‌شود.

برازندگی مربوط به هر ذره در  $x$  و  $P_{best}$  قبلی باهمدیگر مقایسه می‌شوند و براساس نوع مسئله (کمینه سازی یا بیشینه سازی)، اگر برازندگی هر ذره در  $x$  جدید (کمتر یا بیشتر) از معادل آن ذره در  $P_{best}$  باشد، در  $P_{best}$  جدید جایگزین خواهد شد. در نهایت برازندگی ذرات در  $P_{best}$  جدید محاسبه می‌شود و  $G_{best}$  جدید براساس نوع مسئله (کمینه سازی یا بیشینه سازی) از  $P_{best}$  جدید بدست می‌آید.

۷. مراحل ۱ تا ۶ را تا زمان برقرار شدن شرط توقف، تکرار می‌شوند.



شکل ۱: فلوچارت بهینه سازی با الگوریتم ازدحام ذرات



شکل ۲: فلوچارت نحوه عملکرد مدل بهینه سازی

در این مقاله از الگوریتم ازدحام ذرات برای به دست آوردن بهترین مقدار در محدوده تعیین شده برای پارامترهای موجود در مرحله انتخاب ورودی مانند تعیین مقدار ضرایب آستانه موثر بودن ( $TH_1$ ) و آستانه شباهت ( $TH_2$ ) و همچنین تعیین مقدار پارامتر موجود در شبکه عصبی ترکیبی مانند تعداد نورون های لایه پنهان مورد استفاده در هر یک از شبکه عصبی MLP استفاده می‌شود (شکل ۲). تابع هدفی که

$F(i)$  بار پیش بینی شده توسط شبکه عصبی و  $Target(i)$  بردار

هدف متعلق به داده‌های تست می‌باشند.

## ۶- بهینه‌سازی با الگوریتم ازدحام ذرات

الگوریتم ازدحام ذرات اولین بار در سال ۱۹۹۵ توسعه داده شده است. این الگوریتم از دسته الگوریتم‌های بهینه سازی است که بر مبنای تولید تصادفی جمعیت اولیه عمل می‌کنند. در این الگوریتم با الگویی و شبیه سازی رفتار پرواز دسته جمعی (گروهی) پرنندگان یا حرکت دسته جمعی (گروهی) ماهی‌ها بنا نهاده شده است. هر عضو در این گروه توسط بردار سرعت و بردار موقعیت در فضای جستجو تعریف می‌گردد. در هر تکرار زمانی، موقعیت جدید ذرات با توجه به بردار سرعت و بردار موقعیت در فضای جستجو تعریف می‌گردد. در هر تکرار زمانی، موقعیت جدید ذرات با توجه به بردار سرعت فعلی، بهترین موقعیت یافت شده توسط آن ذره و بهترین موقعیت یافت شده توسط بهترین ذره موجود در گروه، به روزرسانی می‌گردد [۲۵]. شکل (۱) فلوچارت الگوریتم بهینه سازی PSO را نشان می‌دهد. الگوریتم PSO شامل مراحل زیر می‌باشد:

۱. موقعیت ذره به طور تصادفی مقداردهی اولیه می‌شود و با توجه به تعداد ذرات ( $Np$ ) و تعداد متغیرهای تصمیم ( $m$ )، ماتریس حاصل توسط  $x$  با  $Np$  سطر و  $m$  ستون به صورت  $x_{Np \times m}$  نشان داده شده است.

۲. همچنین سرعت مربوط به هر ذره به طور تصادفی مقداردهی اولیه می‌شود و با توجه به تعداد ذرات  $Np$  و تعداد متغیرهای تصمیم  $m$ ، ماتریس حاصل توسط  $v$  با  $Np$  سطر و  $m$  ستون به صورت  $v_{Np \times m}$  نشان داده شده است.

۳. بهترین موقعیت پیدا شده توسط هر ذره را  $P_{best}$  و بهترین موقعیت پیدا شده توسط گروهی از ذرات را  $G_{best}$  می‌نامند. در تکرار اول  $P_{best} = x$  می‌باشد و بعد از این که تابع برازندگی مربوط به  $x$  محاسبه می‌شود، مقدار  $G_{best}$  اگر مسئله کمینه سازی باشد  $G_{best} = \min(P_{best})$  و اگر بیشینه سازی باشد  $G_{best} = \max(P_{best})$  خواهد بود که در

این مقاله هدف کمینه سازی تابع برازندگی می‌باشد.

۴. سرعت هر ذره توسط رابطه زیر بروزرسانی می‌شود:

$$v_i^{t+1} = \omega \times v_i^t + c_1 \times r_1 \times (P_{best_i}^t - x_i^t) + c_2 \times r_2 \times (G_{best}^t - x_i^t) \quad (10)$$

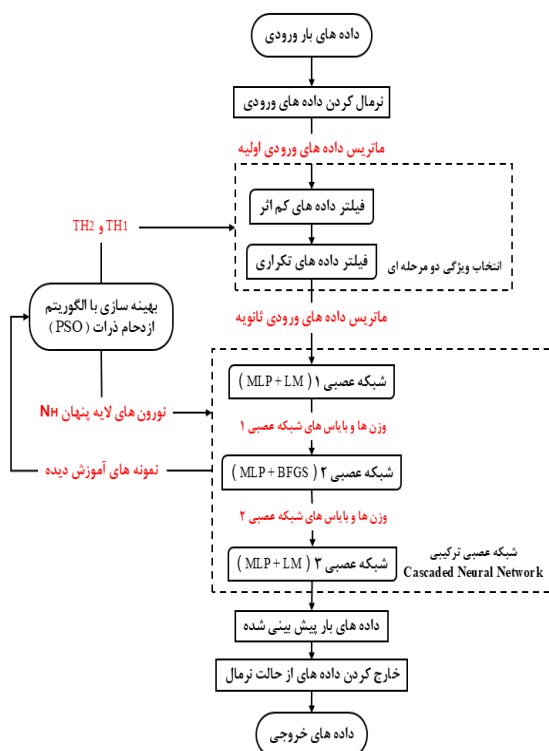
که در رابطه فوق  $t$  تعداد تکرارها ( $t = 1, 2, \dots, t_{max}$ )،  $i = 1, 2, \dots, Np$  در تکرار  $t$ ،  $r_1$  و  $r_2$  اعداد تصادفی بین صفر و یک،  $x_i^t$  موقعیت ذره  $t$ ام در تکرار  $t$ ،  $c_1$  و  $c_2$  توازن بین دانش فردی و اجتماعی را کنترل می‌کنند که در اینجا مقدار هر کدام از آن‌ها ۲ در نظر گرفته شده،  $\omega$  ضریب وزنی است که برای برقراری تعادل بین جستجوی محلی و جستجوی سراسری استفاده می‌شود.

ضریب وزنی ( $\omega$ ) از رابطه زیر محاسبه می‌شود:

۵. وارد شدن نمونه آموزش دیده از قسمت شبکه عصبی به قسمت بهینه ساز از دحام ذارت جهت بهینه کردن پارامترهای قبل تنظیم با کمینه سازی تابع برازندگی MSE می باشد.

۶. برگرداندن خروجی های نرمال شده به حالت اصلی خود یعنی به مقادیر واقعی

۷. محاسبه میزان خطا روش پیشنهادی با استفاده از معیارهای MAPE, RMSE, MSE



شکل ۳: فلوچارت نهایی مدل پیشنهادی در این مقاله

## ۸- بررسی نتایج مدل پیشنهادی

جدول های ۱ تا ۴ مقادیر خطای محاسبه شده برای مدل پیشنهادی در چهار فصل (بهار، تابستان، پاییز و زمستان) را نشان می دهد. همچنین شکل های ۵ تا ۱۲ نمودارهای مقدار بار واقعی و پیش بینی شده را در چهار فصل سال (بهار، تابستان، پاییز و زمستان) با استفاده از روش های LM-MLP-PSO, MLP+BFGS-PSO و CNN-PSO نشان می دهد. بر اساس داده های ارائه شده در جداول و فصل های مختلف، می توان عملکرد سه مدل مختلف را از نظر دقت پیش بینی بررسی و مقایسه کرد. برای تحلیل دقیق، سه شاخص کلیدی ارزیابی عملکرد مدل ها شامل MSE, RMSE و MAPE در نظر گرفته شد. در فصل بهار مدل CNN-PSO با  $MSE = 0.000055$ ,  $MAPE = 0.8654\%$  از سایر مدل های کارایی و دقت پیش بینی بالاتری داشت هر چند میزان خطا RMSE مدل MLP+BFGS-PSO نسبت به سایر مدل ها بهتر بود. داده های جداول نشان می دهد که مدل پیشنهادی این مقاله در فصل تابستان، پاییز و زمستان نیز از لحاظ معیار MSE و MAPE نیز بهتر از دو مدل دیگر می باشد اما همچنان این مدل از لحاظ

قرار است با استفاده از این الگوریتم کمینه شود، همان خطای میانگین مربع (MSE) می باشد، که تعیین مقادیر پارامترهای ذکر شده با توجه به کمترین مقدار به دست آمده از این تابع هدف می باشد [۲۶].

## ۷- بحث در مورد روش پیشنهادی

در این مقاله سعی شده که راه حلی برای مشکل محدودیت آموزش یک شبکه عصبی معمولی و همچنین تنظیم پارامترهای تاثیرگذار در روند کاهش خطای پیش بینی همچون  $TH_1$  و  $TH_2$  در قسمت انتخاب ویژگی دو مرحله ای و تعداد نورون های لایه پنهان  $N_H$  در قسمت شبکه عصبی ترکیبی ارائه شده است.

روش ارائه شده برای حل مشکل محدودیت در آموزش یک شبکه عصبی معمولی، استفاده از شبکه عصبی ترکیبی که حاصل ترکیب سه شبکه عصبی معمولی از نوع پرسپترون چند لایه به صورت متوالی می باشد. همچنین برای حل مشکل تنظیم پارامترهایی همچون  $TH_1$ ،  $TH_2$  و  $N_H$  استفاده از روش بهینه سازی ازدحام ذرات برای پیدا کردن مقدار بهینه هر کدام از این پارامترها و بهبود عملکرد پیش بینی می باشد.

در نتیجه برای پیش بینی بار الکتریکی با استفاده از روش های پیشنهادی در این مقاله (شکل ۳)، بایستی مراحل زیر به ترتیب گذرانده شوند:

۱. دریافت داده های بار الکتریکی و نرمال سازی آن ها در بازه ۰ تا ۱. داده های مورد استفاده در این تحقیق مربوط به بار الکتریکی استان کرمان در سال ۱۳۹۷ بوده که شامل اطلاعات ساعتی برای تمامی ۱۲ ماه سال است و در مجموع ۸۷۶۰ نمونه را در بر می گیرد. این داده ها در شکل ۴ به صورت نمودار زمانی نمایش داده شده اند.

۲. تشکیل ماتریس داده های اولیه و بردار هدف که برای این کار داده های بار الکتریکی ساعتی را برای هر فصل از سال تفکیک کرده و برای هر فصل ۱۵۰۰ نمونه آموزشی ( $TS = 1500$ ) و داده های ۲۰۰ ساعت قبل ( $BS = 200$ ) استفاده شده که از این ۱۵۰۰ نمونه آموزشی، ۵۰۰ نمونه به عنوان داده های تست و بقیه داده ها به عنوان داده های آموزش جهت ورود به مراحل بعدی مدل پیشنهادی، تقسیم بندی شده اند.

۳. انتخاب ویژگی دو مرحله ای، که در مرحله اول فیلتر داده های کم اثر و در مرحله دوم فیلتر داده های تکراری بر روی ماتریس داده های اولیه اعمال می شود، که حاصل این مرحله ایجاد ماتریس داده های ثانویه با داده های موثرتر و غیر تکراری می باشد.

۴. وارد شدن ماتریس داده های ثانویه به شبکه عصبی ترکیبی که متشکل از سه شبکه عصبی متوالی، که هر سه شبکه عصبی از نوع پرسپترون چند لایه با الگوریتم های آموزشی متفاوت و انجام فرآیند آموزش و پیش بینی بار الکتریکی بر روی ماتریس داده های ورودی ثانویه.

۱.۲۹.۵۷.۸۴

جدول ۲: مقادیر خطاهای محاسبه شده برای فصل تابستان

فصل تابستان	MSE		
	MLP+LM+PSO	MLP+BFGS+PSO	CNN+PSO
	۰/۰۰۰۱۴	۰/۰۰۰۳۵	۰/۰۰۰۰۷۵
	RMSE		
	MLP+LM+PSO	MLP+BFGS+PSO	CNN+PSO
	۰/۰۱۲	۰/۰۱۸	۰/۰۰۸
	MAPE %		
	MLP+LM+PSO	MLP+BFGS+PSO	CNN+PSO
	۱/۴۰۳۸	۲/۱۸۰۷	۱/۰۱۱۴
ویژگی های انتخاب شده در این فصل از سال			
۱.۳.۵.۷.۹			

جدول ۳: مقادیر خطاهای محاسبه شده برای فصل پاییز

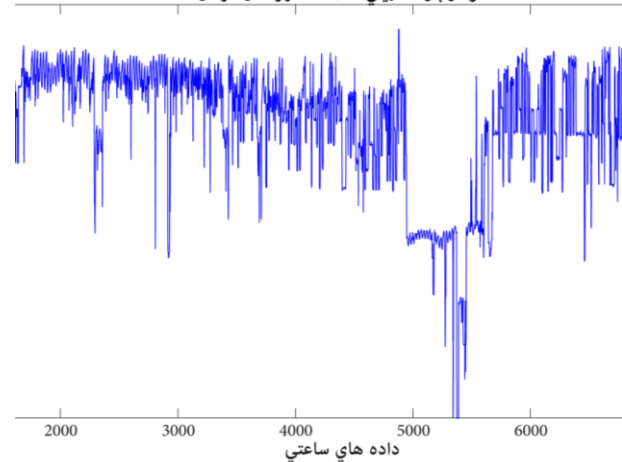
فصل پاییز	MSE		
	MLP+LM+PSO	MLP+BFGS+PSO	CNN+PSO
	۰/۰۰۰۰۴	۰/۰۰۰۰۸	۰/۰۰۰۰۳
	RMSE		
	MLP+LM+PSO	MLP+BFGS+PSO	CNN+PSO
	۰/۰۲۰۵	۰/۰۲۸۴	۰/۰۱۷۳
	MAPE %		
	MLP+LM+PSO	MLP+BFGS+PSO	CNN+PSO
	۲/۷۱۰۶	۳/۷۵۱۳	۲/۳۹۲۸
ویژگی های انتخاب شده در این فصل از سال			
۱.۴۱.۱۲۰.۱۹۲			

جدول ۴: مقادیر خطاهای محاسبه شده برای فصل زمستان

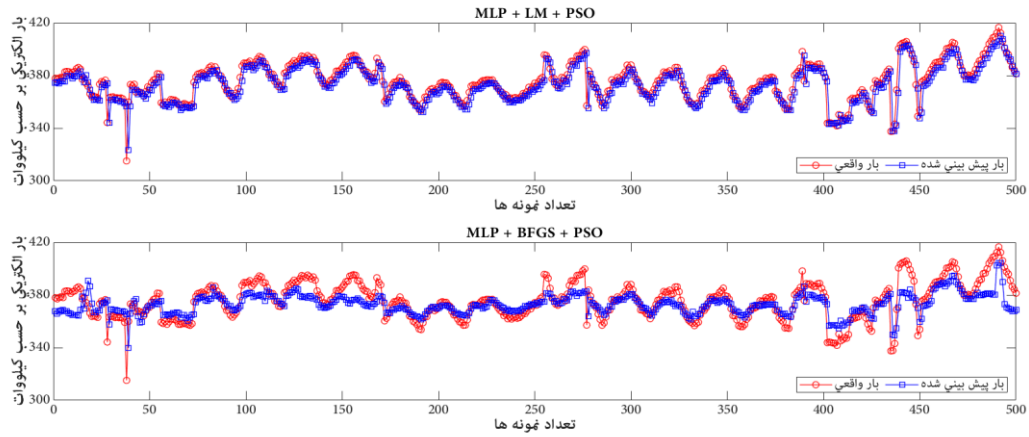
فصل زمستان	MSE		
	MLP+LM+PSO	MLP+BFGS+PSO	CNN+PSO
	۰/۰۰۰۱۱	۰/۰۰۰۱۷	۰/۰۰۰۰۹۲
	RMSE		
	MLP+LM+PSO	MLP+BFGS+PSO	CNN+PSO
	۰/۰۳۴	۰/۰۴۲	۰/۰۳۰
	MAPE %		
	MLP+LM+PSO	MLP+BFGS+PSO	CNN+PSO
	۴/۲۶۶۹	۵/۳۱۲۴	۳/۸۲۴۷
ویژگی های انتخاب شده در این فصل از سال			
۱.۴.۷.۱۰.۱۳.۱۸.۲۳			

RMSE در جایگاه دوم نسبت به مدل MLP-BFGS-PSO می باشد. در این پژوهش، استفاده از الگوریتم PSO نقش کلیدی در افزایش دقت مدل پیش‌بینی بار کوتاه‌مدت ایفا کرده است. از این بهینه‌سازی مرحله‌ی انتخاب ویژگی و بهینه‌سازی تعداد نرون‌ها برای بهینه‌سازی سه مؤلفه کلیدی TH1، TH2، و NH استفاده شده است. نتایج نشان می‌دهد که PSO نه تنها به بهینه‌سازی ساختار مدل کمک کرده، بلکه با حذف داده‌های زلند، انتخاب ویژگی‌های مؤثر و تنظیم هوشمندانه پارامترهای شبکه، موجب کاهش چشمگیر خطای مدل و در نتیجه، دقت نهایی مدل به‌طور معناداری ارتقاء یافته است. این الگوریتم نه تنها توانسته مدل را در برابر داده‌های پیچیده‌تر تطبیق‌پذیرتر کند، بلکه به‌صورت عددی نیز با کاهش محسوس در خطاها، نقش مؤثری در افزایش قابلیت اطمینان پیش‌بینی ایفا کرده است.

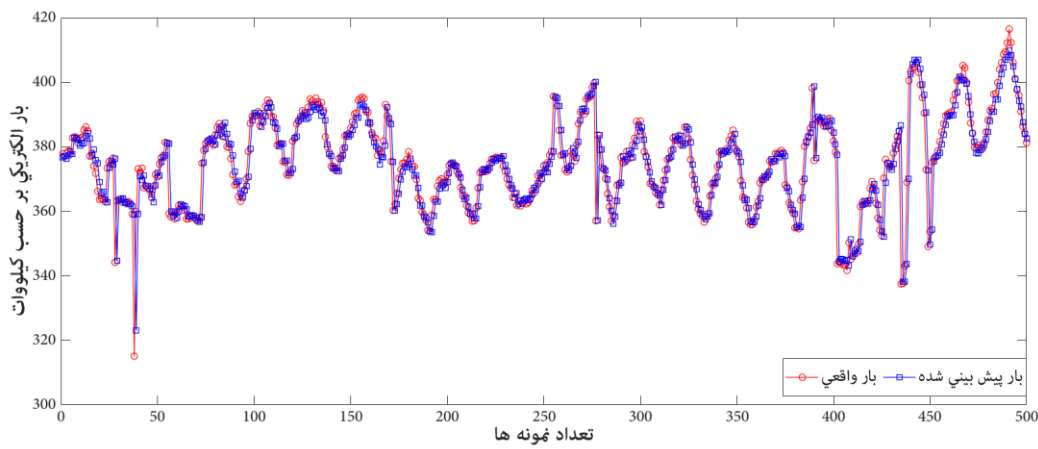
نمودار بار الکتریکی سال ۹۷ شهرستان کرمان

شکل ۴: نمودار داده‌های ساعتی بار الکتریکی سال ۹۷  
جدول ۱: مقادیر خطاهای محاسبه شده برای فصل بهار

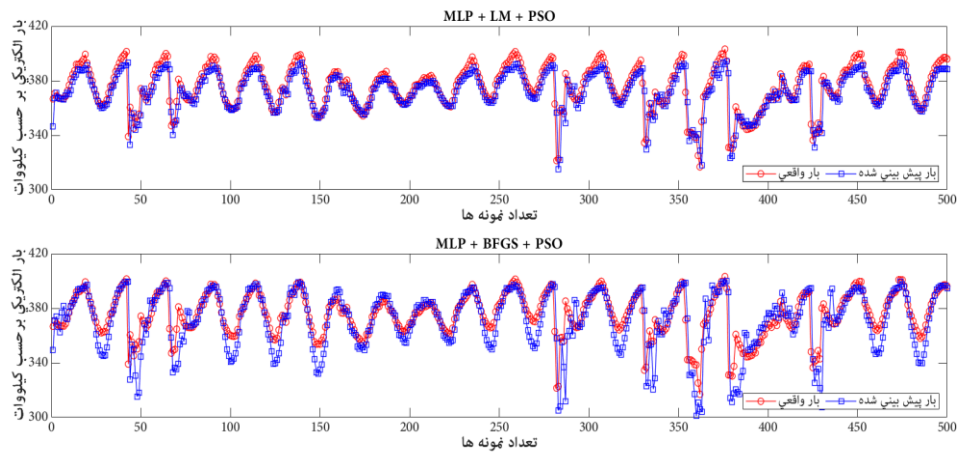
فصل بهار	MSE		
	MLP+LM+PSO	MLP+BFGS+PSO	CNN+PSO
	۰/۰۰۰۰۹۵	۰/۰۰۰۰۲۷	۰/۰۰۰۰۵۵
	RMSE		
	MLP+LM+PSO	MLP+BFGS+PSO	CNN+PSO
	۰/۰۰۹۷	۰/۰۱۶۶	۰/۰۰۷۴
	MAPE %		
	MLP+LM+PSO	MLP+BFGS+PSO	CNN+PSO
	۱/۱۳۵۸	۱/۹۴۰۲	۰/۱۸۶۵۴
ویژگی های انتخاب شده در این فصل از سال			



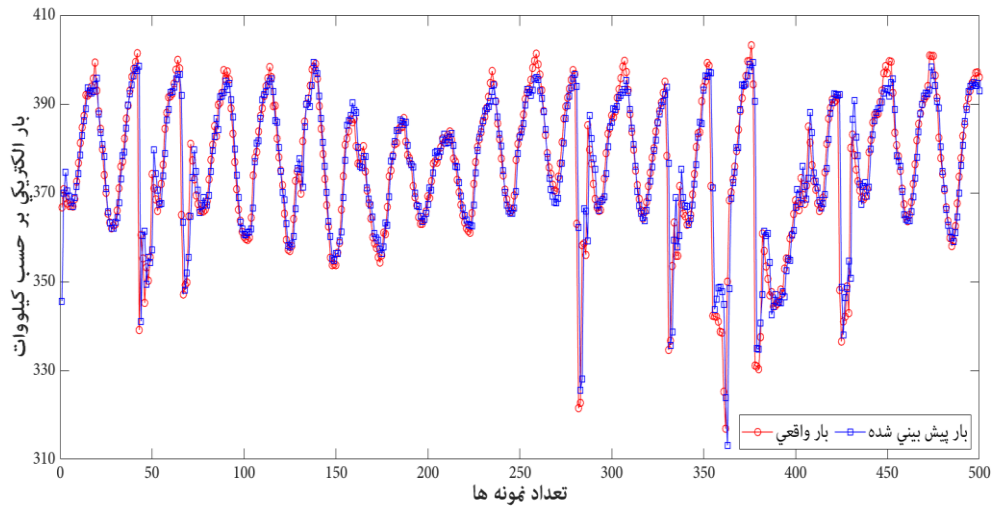
شکل ۵: نمودار مقادیر بار واقعی و پیش بینی شده توسط یک شبکه عصبی MLP+LM و MLP+BFGS با کمک بهینه ساز PSO برای فصل بهار



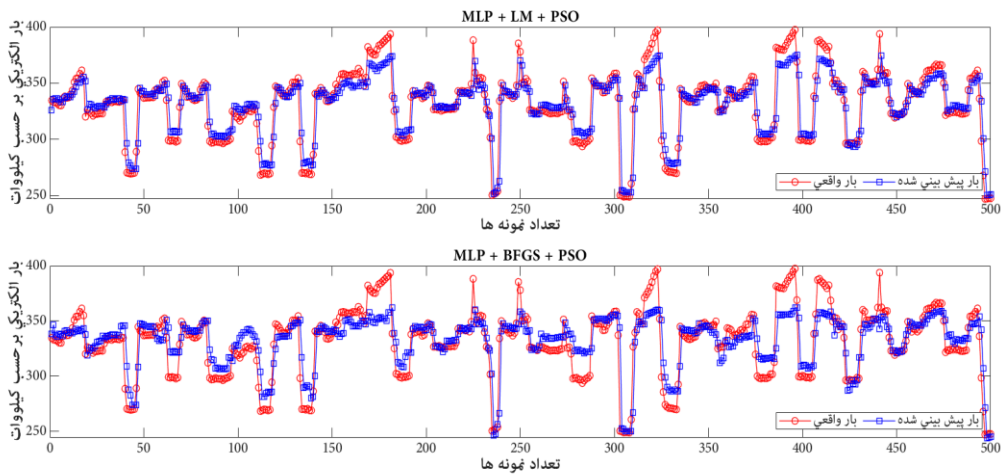
شکل ۶: نمودار مقادیر بار واقعی و پیش بینی شده توسط مدل پیشنهادی با ترکیب شبکه عصبی ترکیبی (CNN) و بهینه-ساز



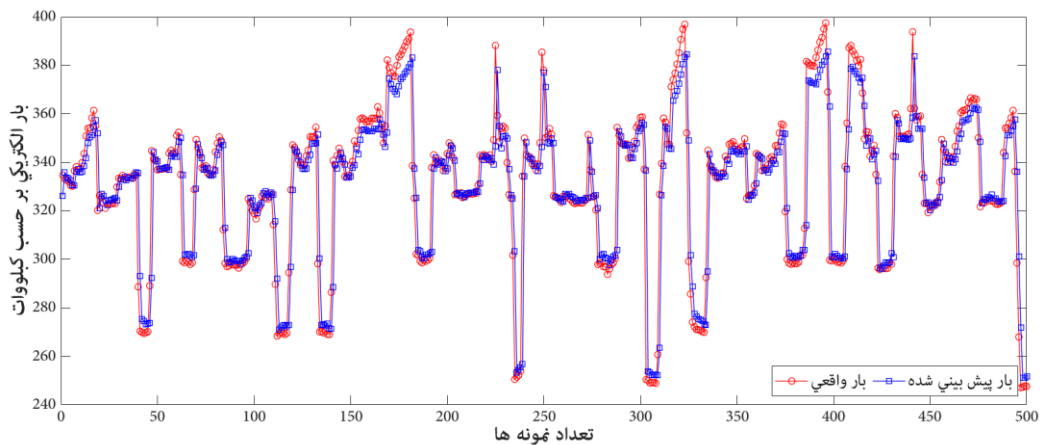
شکل ۷: نمودار مقادیر بار واقعی و پیش بینی شده توسط یک شبکه عصبی MLP+LM و MLP+BFGS با کمک بهینه ساز PSO برای فصل تابستان



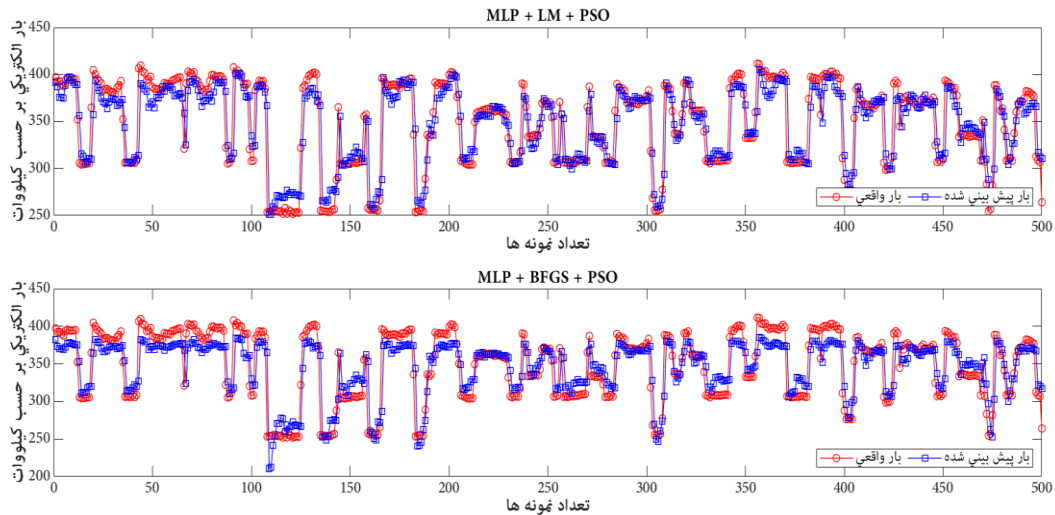
شکل ۸: نمودار مقادیر بار واقعی و پیش‌بینی شده توسط مدل پیشنهادی با ترکیب شبکه عصبی ترکیبی (CNN) و بهینه‌ساز



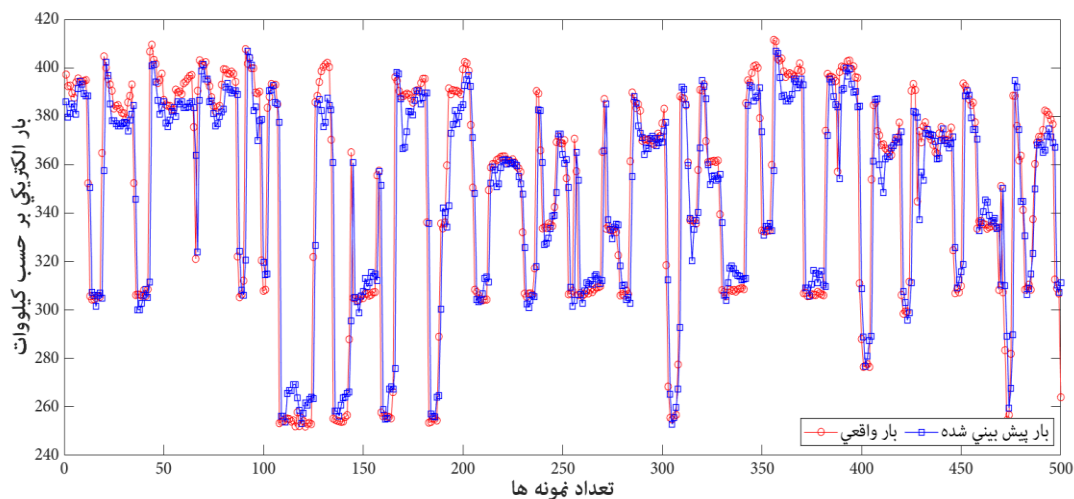
شکل ۹: نمودار مقادیر بار واقعی و پیش‌بینی شده توسط یک شبکه عصبی MLP+LM و MLP+BFGS با کمک بهینه‌ساز PSO برای فصل پاییز



شکل ۱۰: نمودار مقادیر بار واقعی و پیش‌بینی شده توسط مدل پیشنهادی با ترکیب شبکه عصبی ترکیبی (CNN) و بهینه‌ساز PSO برای فصل پاییز



شکل ۱۱: نمودار مقادیر بار واقعی و پیش بینی شده توسط یک شبکه عصبی MLP+LM و MLP+BFGS با کمک بهینه ساز PSO برای فصل زمستان



شکل ۱۲: نمودار مقادیر بار واقعی و پیش بینی شده توسط مدل پیشنهادی با ترکیب شبکه عصبی ترکیبی (CNN) و بهینه ساز PSO برای فصل زمستان

## مراجع

- [1] K. Methaprayoon, W. Lee, P. Didsayabutra, J. Liao, and R. Ross, "Neural network-based short term load forecasting for unit commitment scheduling," in IEEE Technical Conference on Industrial and Commercial Power Systems, 2003., 2003: IEEE, pp. 138-143.
- [2] N. Kamel and Z. Baharudin, "Short term load forecast using Burg autoregressive technique," in 2007 International Conference on Intelligent and Advanced Systems, 2007: IEEE, pp. 912-916.
- [3] S. Ruzic, A. Vuckovic, and N. Nikolic, "Weather sensitive method for short term load forecasting in electric power utility of Serbia," IEEE Transactions on Power Systems, vol. 18, no. 4, pp. 1581-1586, 2003.
- [4] S.-J. Huang and K.-R. Shih, "Short-term load forecasting via ARMA model identification including non-Gaussian process considerations," IEEE Transactions on power systems, vol. 18, no. 2, pp. 673-679, 2003.
- [5] Y. Loewenstern, L. Katzir, and D. Shmilovitz, "Statistical analysis of power systems and application to load forecasting," in 2014 IEEE 28th Convention of Electrical & Electronics Engineers in Israel (IEEEI), 2014: IEEE, pp. 1-5.

## ۹- نتیجه گیری

پیش بینی بار کوتاه مدت در زمینه های مختلف سیستم های قدرت نقش مهمی را ایفا می کند. روش های متفاوت و متنوعی برای پیش بینی کوتاه مدت پارامترهای مختلف سیستم های قدرت مورد استفاده قرار گرفته اند. همچنین گفته شد که بعضی از این روش ها همچون تکنیک های سری زمانی نمی توانند به خوبی پارامترهای غیر خطی سیستم های قدرت را در کوتاه مدت تخمین بزنند، در نتیجه از تکنیک های هوش مصنوعی همچون شبکه عصبی برای تخمین مورد استفاده قرار گرفتند، که این تکنیک ها با تقلید از نحوه تفکر انسان می توانند رابطه بین ورودی و خروجی را یاد بگیرند.

- [26] M. Askari and F. Keynia, "Mid-term electricity load forecasting by a new composite method based on optimal learning MLP algorithm," *IET Generation, Transmission & Distribution*, vol. 14, no. 5, pp. 845-852, 2019.
- [6] M. Q. Raza and A. Khosravi, "A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings," *Renewable and Sustainable Energy Reviews*, vol. 50, pp. 1352-1372, 2015.
- [7] N. Amral, C. Ozveren, and D. King, "Short term load forecasting using multiple linear regression," in *2007 42nd International universities power engineering conference*, 2007: IEEE, pp. 1192-1198.
- [8] F. J. Nogales, J. Contreras, A. J. Conejo, and R. Espinola, "Forecasting next-day electricity prices by time series models," *IEEE Transactions on power systems*, vol. 17, no. 2, pp. 342-348, 2002.
- [9] H. Zareipour, C. A. Cañizares, K. Bhattacharya, and J. Thomson, "Application of public-domain market information to forecast Ontario's wholesale electricity prices," *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp. 1707-1717, 2006.
- [10] J. Contreras, R. Espinola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE transactions on power systems*, vol. 18, no. 3, pp. 1014-1020, 2003.
- [11] R. C. Garcia, J. Contreras, M. Van Akkeren, and J. B. C. Garcia, "A GARCH forecasting model to predict day-ahead electricity prices," *IEEE transactions on power systems*, vol. 20, no. 2, pp. 867-874, 2005.
- [12] J. W. Taylor, "Short-term load forecasting with exponentially weighted methods," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 458-464, 2011.
- [13] D. Infield and D. Hill, "Optimal smoothing for trend removal in short term electricity demand forecasting," *IEEE Transactions on Power Systems*, vol. 13, no. 3, pp. 1115-1120, 1998.
- [14] N. Amjady, "Day-ahead price forecasting of electricity markets by a new fuzzy neural network," *IEEE Transactions on power systems*, vol. 21, no. 2, pp. 887-896, 2006.
- [15] J.-J. Guo and P. B. Luh, "Improving market clearing price prediction by using a committee machine of neural networks," *IEEE Transactions on Power Systems*, vol. 19, no. 4, pp. 1867-1876, 2004.
- [16] Y.-Y. Hong and C.-F. Lee, "A neuro-fuzzy price forecasting approach in deregulated electricity markets," *Electric Power Systems Research*, vol. 73, no. 2, pp. 151-157, 2005.
- [17] M. Mohandes, "Support vector machines for short-term electrical load forecasting," *International Journal of Energy Research*, vol. 26, no. 4, pp. 335-345, 2002.
- [18] C. P. Rodriguez and G. J. Anders, "Energy price forecasting in the Ontario competitive power system market," *IEEE transactions on power systems*, vol. 19, no. 1, pp. 366-374, 2004.
- [19] L. Zhang, P. B. Luh, and K. Kasiviswanathan, "Energy clearing price prediction and confidence interval estimation with cascaded neural networks," *IEEE Transactions on Power Systems*, vol. 18, no. 1, pp. 99-105, 2003.
- [20] G. Bebis and M. Georgiopoulos, "Feed-forward neural networks," *IEEE Potentials*, vol. 13, no. 4, pp. 27-31, 1994.
- [21] محمدی و کی نیا "ارائه یک روش جدید انتخاب ورودی بر اساس دسته‌بندی نیمه نظارتی برای تخمین سری های زمانی," *مدل سازی در مهندسی*, vol. ۱۷, no. ۵۹, pp. ۱۵۳-۱۶۳, زمستان ۱۳۹۸.
- [22] N. Amjady and F. Keynia, "Day-ahead price forecasting of electricity markets by mutual information technique and cascaded neuro-evolutionary algorithm," *IEEE Transactions on Power Systems*, vol. 24, no. 1, pp. 306-318, 2008.
- [23] عابدی نیا و امجدی "پیش‌بینی بار در بازار برق با استفاده از مدل هیبرید شبکه عصبی و الگوریتم بهبود یافته جستجوی گرانشی," *مدل سازی در مهندسی*, vol. ۱۳, no. ۴۰, pp. ۶۹-۷۷, ۱۳۹۸.
- [24] F. Keynia, "A new feature selection algorithm and composite neural network for electricity price forecasting," *Engineering Applications of Artificial Intelligence*, vol. 25, no. 8, pp. 1687-1697, 2012.
- [25] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, 1995: IEEE, pp. 39-43.