



## پیش‌بینی قیمت مس به کمک شبکه عصبی مصنوعی

رضا رسام<sup>۱</sup>، محمدحسین بصیری<sup>۲</sup>

۱ دانشجوی کارشناسی ارشد، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس تهران، ایران

آدرس پست الکترونیکی ([rezarasam@modares.ac.ir](mailto:rezarasam@modares.ac.ir))

۲ استادیار، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس تهران، ایران

آدرس پست الکترونیکی ([mhbasiri@modares.ac.ir](mailto:mhbasiri@modares.ac.ir))

### چکیده

بدلیل نقش پررنگ مس در بازارهای بورس فلزات، شناخت از قیمت آینده این فلز از اهمیت بالایی برخوردار است. در این مقاله ابتدا پارامترهای موثر بر قیمت مس بررسی می‌شوند. سپس روش تحلیل مولفه‌های اساسی جهت کاهش ابعاد پارامترهای ورودی مورد استفاده قرار می‌گیرد و در نهایت شبکه عصبی پس‌انتشار پیشخور جهت پیش‌بینی قیمت مس ارائه می‌شود. نتایج این مقاله برتری شبکه عصبی سه‌لایه با ساختار ۱-۱۰-۸-۶ آموزش یافته توسط الگوریتم LM جهت پیش‌بینی قیمت مس نسبت به سایر ساختارها و الگوریتم‌ها، را نشان می‌دهد. RMSE مدل بهینه برابر با ۲۴۳/۳۴ بدست آمده است.

کلمات کلیدی: مس، پیش‌بینی، تحلیل مولفه‌های اساسی، شبکه عصبی پیشخور پس‌انتشار



## Copper price forecasting with the help of artificial neural network

Reza, Rasam<sup>1</sup>; MohammadHosein, Basiri<sup>2</sup>

1-Msc student, Department of engineering, University of Tarbiat Modares, Tehran, Iran

(Email: [rezarasam@modares.ac.ir](mailto:rezarasam@modares.ac.ir))

2- Assistant Professor, Department of engineering, University of Tarbiat Modares, Tehran, Iran

(Email: [mhbasiri@modares.ac.ir](mailto:mhbasiri@modares.ac.ir))

### Abstract

Because of the copper role in the metals exchange markets, knowledge of the future price of the metal is very important. In this paper, the parameters affecting the price of copper are investigated. Then the principal components analysis to reduce the size of the input parameters used and finally feedforward back propagation neural network to forecast the price of copper is provided. The paper's results indicate that three-layer neural network with 6-8-10-1 structure and LM training algorithm to forecast copper prices is better than other structures and algorithms. And RMSE for this model is 243.34.

**Keywords:** *Copper, Forecasting, Principal components analysis, Neural networks feedforward BP*

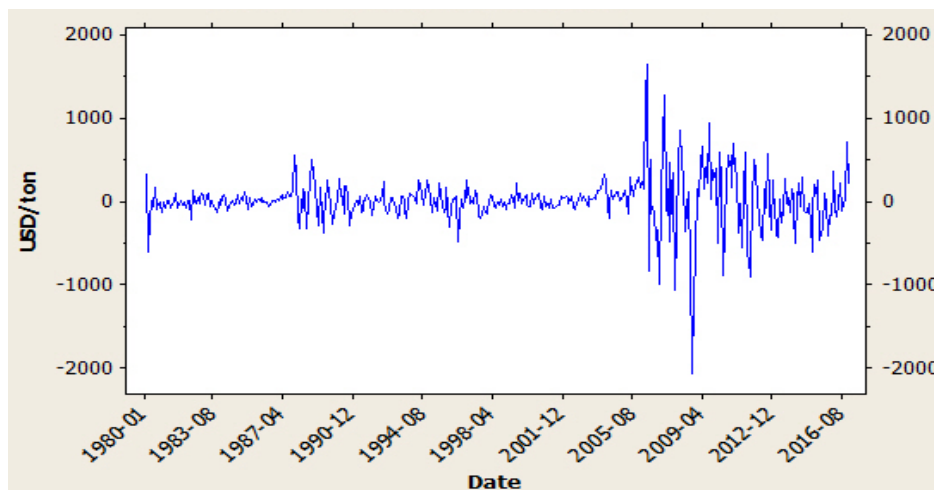
## ۱- مقدمه

پیش بینی هنر و علم شناخت و خبر دادن از حوادث آینده می باشد. امروزه پیش بینی و شناخت متغیرهای اقتصادی، از اهمیت بالایی برای سیاست گذاران در تنظیم روابط اقتصادی برخوردار است [۱].

یکی از هدف های اصلی تجزیه و تحلیل های اقتصادی، پیش بینی صحیح و دقیق متغیرهای اقتصادی می باشد که سیاست گذاران را جهت تصمیمات صحیح و متناسب با مقادیر پیش بینی شده یاری می کند [۲]. واضح است که هرچه پیش بینی دقیق تر باشد تصمیمات متاثر از آن نیز به شکل مناسب تری گرفته خواهد شد.

مس بدلیل ویژگی هایی چون چکش خواری، مقاومت در برابر خوردگی، رسانای خوب الکتریکی و حرارتی و ساخت آلیاژهای مختلف مانند برنج و برنز، کاربردهای فراوانی دارد و از این رو در بازارهای معاملات جهانی مورد معامله قرار می گیرد. همچنین تولید و استفاده از آن ضروری شکل ۱- نوسان قیمت مس نسبت به ماه قبل بوده و فاکتور مهمی در تضمین توسعه پایداری یک جامعه است. بازار مس بیش از هر عامل دیگری وابسته به نقضای جهانی می باشد [۳]. از این رو پیشرفت تکنولوژی، وضعیت صنعتی کشورها و همچنین شرایط اقتصادی کشورهای بزرگ مصرف کننده می تواند بر قیمت مس تاثیر زیادی داشته باشد. به همین دلیل کارشناسان مس را نوعی داماسنج اقتصاد بین الملل می دانند [۳].

با توجه به شکل ۱ قیمت مس در سال های اخیر همانند سایر فلزات نوسانات بالایی داشته و در شناخت قیمت بازار آینده این فلز نیز عدم قطعیت زیادی وجود دارد. با شناخت از قیمت آینده مس، مصرف کنندگان و تولید کنندگان قادر به مدیریت ریسک بوده و مصرف کنندگان می توانند قیمت خرید امن و تولید کنندگان قیمت فروش امنی داشته باشند همچنین سرمایه گذاران می توانند با شناخت از بازار، ریسک معاملات را کاهش دهند [۷].



## ۲- پیشینه تحقیق

اسمیت و براکر [۸ و ۹] در سال های ۱۹۹۹ و ۲۰۰۳ مدل های GARCH و EGARCH را جهت پیش بینی نوسانات قیمت مس در آینده مورد استفاده قرار داده اند و برای توصیف مناسب تر رفتار نوسانی آینده قیمت مس راهکارهایی پیشنهاد می دهند. در سال ۲۰۰۴، بومن و همکاران [۱۰] قیمت آینده پانزده کالای مختلف از جمله مس را توسط مدل خودرگرسیو ARIMA پیش بینی کرده اند. در سال ۲۰۰۵، خدایاری و رحیمی [۶] قیمت مس در سال های ۱۹۹۵ تا ۲۰۰۴ را به روش های

میانگین متحرک، میانگین متحرک وزنی، هموارسازی نمایی و هموارسازی نمایی تعدیل شده برآورد کرده اند. در سال ۲۰۰۷، فرتی و گیلبرت [۱۱] روش FIGARCH دوتا شونده را برای قیمت مس و آلومینیوم پیشنهاد داده اند. واتکینز و مک آلیر [۱۲] در سال ۲۰۰۸ از روش AR(1)-GARCH(1,1) جهت برآورد و پیش بینی فرآیندهای نوسانی قیمت آینده روزانه برای دو فلز غیر آهنی یعنی آلومینیوم و مس استفاده کرده اند. هاموده و همکاران [۱۳] در سال ۲۰۰۸ تاثیر نرخ مبادله دلار به یورو را بر نوسان های کالاهایی چون آلومینیوم، مس، طلا و نفت توسط مدل های GARCH بررسی کرده اند. در سال ۲۰۱۰، چن [۱۴] در پژوهش خود، جهت بررسی وضعیت قیمت ها و خصوصیات ۲۱ فلز پایه از سری های زمانی بین سال های ۱۹۰۰ تا ۲۰۰۷ استفاده کرده است. کرتازر و اترویک [۱۵] در سال ۲۰۱۰ مدلی برای برآورد قیمت قراردادهای آینده بلندمدت مس و نقره با استفاده از قیمت قراردادهای آینده نفت ارائه داده اند. بیولت [۱۶] در سال ۲۰۱۰ مدل های آماری EGARCH، GARCH و TGARCH را جهت پیش بینی قیمت فلزات آلومینیوم، مس، سرب، قلع، نیکل و روی استفاده کرده است. گارسیاچیکو و مونترو [۱۷] در سال ۲۰۱۱ به بررسی خواص مدل های سوئیچینگ مارکوف برای توصیف رفتار قیمت مس پرداخته اند. لاگوس [۱۸] در سال ۲۰۱۲، شاخصی به نام شاخص نوسان ها معرفی کرده است. در این مطالعه این شاخص برای آهن، مس و چند فلز دیگر محاسبه می شود. اوسار و گاس [۱۹] در سال ۲۰۱۳ در بازار مس بورس فلزات لندن (LME) یک مدل را به نمایندگی از فعالیت های دلان و مصرف کنندگان توسعه داده اند. روتلج و همکاران [۲۰] در سال ۲۰۱۳ رابطه بین قیمت آینده مس در بازارهای جهانی و انتقال اطلاعات و تاثیرات این بازارها بر یکدیگر را بررسی کرده اند. لاشراز و همکاران [۷] در سال ۲۰۱۵ عملکرد ARIMA و دو روش شبکه عصبی MLP و Elman را جهت پیش بینی قیمت مس در بازار جهانی COMEX مورد ارزیابی قرار داده اند. بونسیک و مورتو [۲۱] در سال ۲۰۱۵ با استفاده از داده های LME و استفاده از مدل های DMA و DMS قیمت مس را پیش بینی کرده اند.

هدف اصلی این تحقیق ارائه مدلی مناسب جهت پیش بینی قیمت ماهانه مس می باشد. در بخش بعدی پارامترهای موثر بر قیمت مس معرفی و بررسی می شوند. در بخش ۴، روش های نرمال سازی معرفی شده و همچنین روش های کاهش ابعاد مورد بررسی قرار می گیرند. در بخش ۵، شبکه عصبی مصنوعی<sup>۱</sup> و همچنین روش های ارزیابی شبکه عصبی معرفی می شوند. در بخش ۶ مدل های مختلف با ساختارهای متفاوت بررسی شده و بهترین مدل ها با یکدیگر مقایسه و در نهایت بهترین معماری شبکه معرفی می شود. و سرانجام در بخش ۷، نتیجه گیری ارائه می گردد.

### ۳- پارامترهای موثر بر قیمت مس

عوامل زیادی بر قیمت فلزات از جمله قیمت مس تاثیرگذار هستند. شناسایی این عوامل جهت شناخت قیمت آینده مس از اهمیت بالایی برخوردار است عواملی چون عرضه و تقاضای مس، در بازارهای جهانی، قیمت نفت خام، طلا و آلومینیوم، شاخص قیمت مصرف کننده<sup>۲</sup>، تبدیل ارز یورو به دلار، تولید ناخالص داخلی<sup>۳</sup> و... بر قیمت مس تاثیرگذار می باشند. جدول ۱ همبستگی پارامترهای موثر شناسایی شده بر قیمت مس را نشان می دهد. لازم بذکر است که هرچه همبستگی به عدد یک و یا منفی یک نزدیک تر باشد، تغییرات آن پارامتر بر قیمت مس تاثیر بیشتری دارد. هدف اصلی در این مقاله استفاده از حداکثر این پارامترها جهت پیش بینی قیمت مس با کمک روش کاهش ابعاد تحلیل مولفه های اساسی<sup>۴</sup> می باشد.

جدول ۱- همبستگی پارامترها با قیمت مس

1 Artificial neural network (ANN)

2 Consumer price index (CPI)

3 Gross domestic product (GDP)

4 Principal components analysis (PCA)



همبستگی<sup>۱</sup> با قیمت

پارامتر موثر

مس	
۰/۹۱۸	قیمت نفت خام
۰/۸۶۷	قیمت طلا
۰/۸۱۸	قیمت آلومینیوم
۰/۸۸۱	تولید ناخالص داخلی (GDP) جهانی
۰/۷۷۵	جمعیت جهان
۰/۶۸	شاخص قیمت مصرف کننده (CPI) کشور چین
۰/۶۳۲	تبدیل ارز یورو به دلار (EURO/USD)
-۰/۵۶۸	نرخ بهره چین و آمریکا
۰/۲	شاخص مدیران خرید (PMI) جهانی
۰/۰۱۵	نرخ تورم جهان

جهت انجام پیش‌بینی دقیق و جلوگیری از ساخت مدل پیچیده و همچنین استفاده از تمام پارامترها، برای کاهش پارامترهای ورودی شبکه از روش PCA استفاده می‌شود. در مدل‌های با ساختار پیچیده، امکان وقوع بیش‌برازش<sup>۲</sup> محتمل می‌باشد، بدین معنی که شبکه داده‌های آموزش را حفظ می‌کند و در بخش آزمون توانایی پیش‌بینی را نداشته و خطای زیادی بدست خواهد آمد.

#### ۴- تحلیل و تفسیر داده‌ها

##### ۴-۱- نرمال‌سازی

بدلیل اینکه داده‌ها پراکندگی‌های مختلفی دارند و هرکدام در بازه‌های متفاوتی قرار گرفته‌اند و جهت بهبود عملکرد شبکه عصبی و همچنین نزدیک بودن مقادیر وزن‌ها و بایاس‌های شبکه به یکدیگر جهت جلوگیری از پیچیدگی شبکه، تمام داده‌های ورودی را تبدیل به داده‌هایی در بازه مشخص می‌کنند؛ به این عمل نرمال‌سازی داده‌ها می‌گویند. نرمال‌سازی انواع مختلفی دارد که در ادامه بررسی می‌شوند.

##### ۴-۱-۱- روش خطی

در این روش داده‌ها بصورت خطی در بازه بین صفر و یک قرار می‌گیرند. در این روش بزرگترین نمونه به عنوان حداکثر مقدار یک و کوچکترین نمونه به عنوان حداقل مقدار صفر در نظر گرفته می‌شود و سایر نمونه‌ها بین این دو مقدار قرار می‌گیرند.

$$x \longleftrightarrow \hat{x}$$

(معادله ۱)

<sup>1</sup> Correlation

<sup>2</sup> Over fitting

<sup>۳</sup> ضرایب ثابتی هستند که هر گره در شبکه عصبی بخود اختصاص می‌دهد



$$x \in [x_{min}, x_{max}] \implies \hat{x} \in [0,1]$$

$$\hat{x} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

لازم بذکر است که میزان افزایش و یا کاهش  $x$  و  $\hat{x}$  متناسب می باشد.

روش نرمال سازی خطی را می توان با استفاده از معادله ۲ به بازه  $[-1,1]$  تبدیل نمود.

$$\hat{x} = 2 \times \left( \frac{x - x_{min}}{x_{max} - x_{min}} \right) - 1 \quad \hat{x} \in [-1,1] \quad (\text{معادله ۲})$$

### ۲-۱-۴ روش سفیدسازی<sup>۱</sup>

این روش داده ها را به داده های نرمال استاندارد با میانگین صفر و انحراف معیار یک تبدیل می کند.

$$x \iff \hat{x} \begin{cases} \mu = 0 \\ \sigma = 1 \end{cases} \quad \text{یک متغیر}$$

$$\begin{cases} \mu = \text{mean}(x) \\ \sigma = \text{std}(x) \end{cases} \implies \hat{x} = \frac{x - \mu}{\sigma} \quad (\text{معادله ۳})$$

چندمتغیر: در این حالت زمانی که داده ها همبستگی نداشته باشند نرمال استاندارد داده ها با استفاده از معادله ۴ محاسبه می شود.

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \implies \vec{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}, \quad \sigma = \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_n \end{bmatrix} \quad \vec{\hat{x}} = \begin{bmatrix} \frac{x_1 - \mu_1}{\sigma_1} \\ \frac{x_2 - \mu_2}{\sigma_2} \\ \vdots \\ \frac{x_n - \mu_n}{\sigma_n} \end{bmatrix} \quad (\text{معادله ۴})$$

در غیر اینصورت بایستی این وابستگی را لحاظ کنیم.

$$\hat{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \xrightarrow{\text{cov}} \Sigma \in R^{n \times n} \quad \Sigma \geq 0$$

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix}$$

در صورتی که مقادیر غیر از قطر اصلی صفر باشند همبستگی بین

پارامترها وجود ندارد در غیر اینصورت بایستی این همبستگی

لحاظ شود

$$\hat{x} \begin{cases} \vec{\mu} : \vec{0} \\ \vec{\sigma} : I \end{cases}$$

(معادله ۵)

<sup>1</sup> whitening



$$\hat{\vec{x}} = Q^{-1}(\vec{x} - \vec{\mu})$$

$$Q^T Q = \Sigma$$

در معادله ۵، Q را جذر چولگی می نامند.

### ۳-۱-۴- نگاهت غیر خطی

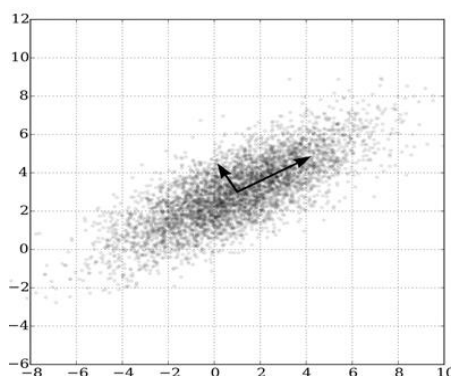
این روش نرمال سازی زمانی که پارامترها تناسب خاصی نسبت به یکدیگر دارند و مقادیر بزرگ هستند استفاده می شود. به عنوان مثال در صورتی که نمونه های یک پارامتر حول  $10^3$  و نمونه های پارامتر دیگر حول  $10^4$  باشد، می توان داده ها را به لگاریتم تبدیل کرد.

$$\begin{cases} x_1 & \propto 10^3 & \xrightarrow{\log_{10}} & \hat{x}_1 & \propto 3 \\ x_2 & \propto 10^4 & & \hat{x}_2 & \propto 4 \end{cases}$$

### ۲-۴- روش تحلیل مولفه اساسی (PCA)

جهت استفاده از تمام داده های موثر و مقابله با عدم بیش برآزش، می توان از روش تحلیل مولفه اساسی یا PCA استفاده نمود. روش PCA مولفه ها یا داده های ورودی را به داده هایی تبدیل می کند که همبستگی بین آن ها صفر می باشد و همچنین می توان داده های آخر را بدلیل همبستگی خیلی پایین با داده هدف<sup>۱</sup>، حذف نمود و به همین جهت تعداد پارامترهای ورودی را با در نظر گرفتن عدم کاهش دقت مدل، کاهش داد.

اساس کار روش PCA تغییر مولفه ها در جهتی که بیشترین همبستگی یا انحراف معیار را نسبت به داده هدف نشان می دهد، می باشد. بدین ترتیب ابتدا اولین مولفه را در راستای بیشترین واریانس یا پراکندگی انتخاب کرده و سپس سایر مولفه ها بترتیب عمود بر مولفه قبلی و در جهت بیشترین واریانس انتخاب می شوند تا به تعداد پارامترهای اولیه مولفه های جدید تولید شوند. شکل ۲ یک مثال است که در آن دو پارامتر ورودی و یک پارامتر خروجی برای فهم و درک بهتر موضوع و در فضای دوبعدی نشان داده شده است.



شکل ۲- نمونه داده هایی در فضای دوبعدی و تغییر مولفه ها در جهت بیشترین پراکندگی [۲۶]

با توجه به شکل بالا همانطور که پیشتر نیز اشاره شد مولفه اول در جهت بیشترین پراکندگی انتخاب می شود. از این رو می توان داده های آخر که کمترین همبستگی را با داده های هدف دارند، از مدل حذف کنیم. از دیگر خصوصیات استفاده از PCA،

<sup>1</sup> Target



می توان به استفاده از پارامترهای موثر بیشتر در مدل بدون پیچیده شدن ساختار شبکه آن، اشاره کرد. در این تحقیق، پارامترها با استفاده از این روش کاهش بعد می یابند.

## ۵- مبانی شبکه عصبی مصنوعی

### ۵-۱- شبکه عصبی مصنوعی

روزنبلات<sup>۱</sup> در دهه ۱۹۵۰ و ویدرو و هاف<sup>۲</sup> در دهه ۱۹۶۰ پایه و اساس شبکه عصبی را تشکیل داده اند و اولین بار توسط روزنبلات شبکه عصبی پرسپترون ارائه شد [۲]. در اواسط دهه ۱۹۸۰ شبکه عصبی جایگاه ویژه در بین محققان یافت و بعنوان ابزار محاسباتی قوی جهت حل مسائل پیش بینی شناخته شد [۲۲].

شبکه عصبی مصنوعی یک ابزار وسیع، قوی و منعطف جهت آنالیز داده ها و مدل کردن معادلات غیرخطی با دقت بسیار بالا می باشد. ویژگی اساسی مدل شبکه عصبی مستقل بودن آن از فرض های آماری متغیرها، استفاده از تکنیک های موازی و غیر خطی آن ها می باشد. با وجود تنوع بالا در شبکه های عصبی، تمامی آن ها یک ساختار مشابه دارند [۲۳]. شبکه های عصبی مصنوعی عملکرد بسیار خوبی در برخورد با روابط غیر خطی در ورودی ها دارند [۲۲]. شبکه های عصبی پس انتشار (BP) معمول ترین شبکه برآورد کننده چندلایه می باشند که در ۸۰ درصد مواقع مورد استفاده قرار گرفته است [۲۴]. این شبکه ها در سال ۱۹۸۶ توسط روملهارت<sup>۳</sup> و همکاران پیشنهاد و ثابت شده است [۲۵] که بطور موفقیت آمیزی در شاخه های مختلف علوم مورد استفاده قرار می گیرند.

شبکه عصبی مصنوعی از مغز انسان الهام گرفته شده است اما شبکه های عصبی مصنوعی فعلی هرگز به پیچیدگی مغز انسان نیستند. دو شباهت بین شبکه عصبی مغز انسان و شبکه عصبی مصنوعی وجود دارد؛ اول اینکه ساختار هردو از یک ابزار محاسباتی ساده با به هم پیوستگی بسیار بالا تشکیل شده اند و شباهت دوم در این است که در هر دو مورد اتصالات بین نرون ها تابع شبکه را تعیین می کند [۴]. شبکه های عصبی ابزار بسیار مناسبی برای برازش توابع و کشف الگوها بین پارامترها می باشد. یک شبکه عصبی ساده قادر به برازش هر تابع عملیاتی می باشد [۴] از این رو ابزار مناسبی جهت پیش بینی نیز هستند.

مهمترین و پرکاربردترین شبکه عصبی، شبکه عصبی پرسپترون چندلایه می باشد. ساختار این شبکه شامل یک لایه ورودی<sup>۴</sup>، که دربرگیرنده پارامترهای ورودی بوده، یک یا چند لایه میانی یا مخفی<sup>۵</sup> که عملیات پردازش در این قسمت انجام می شود و همچنین لایه خروجی<sup>۶</sup> که شامل پارامتر یا پارامترهای خروجی می باشد؛ این لایه، لایه پنهان قبلی را توسط توابع انتقال<sup>۷</sup> به خروجی یا خروجی ها مرتبط می کند. هر لایه شامل یک یا چند نرون<sup>۸</sup> می باشد که تعداد نرون های لایه خروجی برابر با تعداد خروجی ها بوده و نرون های لایه های میانی قابل تنظیم و تغییر هستند. هرکدام از این نرون ها به تمام نرون های لایه قبل و بعد لایه مذکور مرتبط می باشند. شکل ۳ شمای کلی یک ساختار شبکه عصبی را نشان می دهد.

1 Rusenblat

2 Widrow and Hoff

3 Rumelhart

4 Input layer

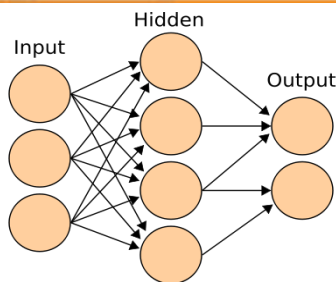
5 Hidden layer(s)

6 Output layer

7 Transfer function

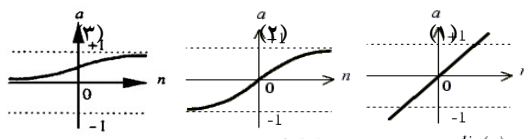
8 Neuron





شکل ۳- نمونه‌ای از ساختار شبکه عصبی دو لایه [۲۶]

در هر نرون عملیات محاسباتی استفاده از توابعی موسوم به توابع انتقال انجام می‌شود. این توابع بردارهای ورودی به نرون را توسط یک تابع دریافت کرده و به هر کدام وزنی مشخص داده و همچنین هر نرون ضریب ثابتی به نام بایاس نیز در اختیار می‌گیرد و سپس خروجی تابع به عنوان خروجی نرون در نظر گرفته می‌شود. توابع انتقال مختلفی وجود دارد، مهمترین آن‌ها که در شبکه‌های عصبی پیشخور مورد استفاده قرار می‌گیرند شامل تانزانت سیگموئید، لگاریتم سیگموئید و تابع خطی می‌باشند.



شکل ۴- توابع انتقال: (۱) تابع انتقال خطی (۲) تابع انتقال تانزانت سیگموئید (۳) تابع انتقال لگاریتم سیگموئید [۴]

شکل ۴ توابع انتقال پر کاربرد در شبکه عصبی پیشخور را نشان می‌دهد. تا این مرحله تنها پوسته شبکه عصبی ساخته می‌شود و مهمترین بخش آن آموزش شبکه بوده که بر عهده الگوریتم‌های آموزش می‌باشد که پردازش اصلی و کیفیت شبکه را بررسی می‌کنند. الگوریتم‌های پس‌انتشار<sup>۱</sup> مناسب‌ترین الگوریتم‌ها جهت آموزش شبکه‌های عصبی پرسپترون می‌باشند.

## ۵-۲- روش‌های ارزیابی عملکرد شبکه

جهت ارزیابی عملکرد شبکه از مقادیر آماری مختلفی استفاده می‌شود که مهمترین‌های آن‌ها در ادامه بررسی می‌شوند.

### ۵-۲-۱- ضریب تعیین

این معیار تعیین کننده میزان پراکندگی داده‌ها توسط مدل می‌باشد و میزان انطباق مدل را نشان می‌دهد [۵]. این ضریب با  $R^2$  نشان داده می‌شود و با استفاده از معادله ۶ محاسبه می‌شود.

$$R^2 = \frac{SSR}{SST} = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(\hat{y}_i - \bar{y})^2 + \sum(y_i - \hat{y}_i)^2} \quad (\text{معادله ۶})$$

<sup>1</sup> Back Propagation Algorithm

### ۲-۲-۵- مجذور میانگین مجموع مربعات خطا<sup>۱</sup>

این معیار برابر با جذر ریشه دوم MSE می باشد. این معیار جهت ارزیابی عملکرد شبکه عصبی کاربرد بالایی دارد.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n}} \quad (\text{معادله ۷})$$

### ۳-۲-۵- میانگین قدرمطلق خطا<sup>۲</sup>

این معیار متوسط اختلاف مقدار واقعی و مقدار تخمین زده شده را نشان می دهد.

$$MAE = \frac{\sum |y_i - \hat{y}_i|}{n} \quad (\text{معادله ۸})$$

### ۴-۲-۵- میانگین قدرمطلق درصد خطا<sup>۳</sup>

این معیار همانند RMSE محبوبیت زیادی دارد و پر استفاده ترین معیار بدون واحد می باشد و تنها در داده های با مقیاس نسبی (وزن، طول، قیمت) کاربرد دارد [۵].

$$MAPE = \frac{1}{n} \sum \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \quad (\text{معادله ۹})$$

این معیار بدون واحد زمانی، از محبوبیت بالایی برخوردار است و برای مقایسه چند سری زمانی مختلف با مقیاس داده های متفاوت مورد استفاده قرار می گیرد. [۵].

## ۶- مدل

در این مقاله، قیمت مس توسط روش شبکه عصبی مصنوعی با الگوریتم های مختلف پیش بینی می شود. از داده های قیمت مس و پارامترهای موثر از سال ۱۹۹۰ تا سال ۲۰۱۶ بصورت ماهیانه یعنی ۳۲۴ داده، استفاده می شود. داده ها به سه دسته داده های آموزش<sup>۴</sup>، ارزیابی<sup>۵</sup> و آزمایش<sup>۶</sup> تقسیم شده اند که این سه دسته داده، بصورت تصادفی از بین داده ها انتخاب می شوند.

همانطور که در بخش ۳ بیان شد پارامترهای مختلفی بر قیمت مس تاثیر گذارند از این رو ۹ پارامتر موثر جهت استفاده بعنوان ورودی شبکه در نظر گرفته شده است و همچنین سه تاخیر زمانی<sup>۷</sup> قیمت مس نیز بعنوان پارامترهای ورودی مورد استفاده قرار می گیرند. سپس داده ها در بازه [۱ و -۱] نرمال می شوند. بدلیل اینکه ۱۲ پارامتر ورودی، برای شبکه در نظر گرفته شده است و از آنجایی که جهت جلوگیری از بیش برآزش شبکه، بایستی شبکه ای با پیچیدگی کم در نظر گرفته شود، روش PCA جهت

1 Root Mean Square Error (RMSE)

2 Mean Absolute Error (MAE)

3 Mean Absolute Percentage Error (MAPE)

4 Training Data

5 Validation Data

6 Test Data

7 Time Delay

کاهش ورودی ها استفاده می شود. با توجه به مطالب بیان شده در بخش ۴-۲، بدلیل تاثیر پایین پارامترهای روبه آخر، می توان از آن ها چشم پوشی کرد. در نتیجه در این مقاله ۶ پارامتر ابتدایی PCA با بیشترین همبستگی، بعنوان ورودی های شبکه در نظر گرفته می شوند.

در این بخش، برای اینکه پارامترهای شبکه به شکل مناسب تری انتخاب شوند، پارامترهای مختلف بصورت جداگانه بررسی شده و بهترین پارامترها توسط روش های ارزیابی معرفی شده در بخش ۵-۲، تعیین می شوند و در نهایت بهترین و مناسب ترین ساختار شبکه معرفی می گردد.

### ۱-۶- انتخاب مناسب ترین توابع انتقال

در این بخش با استفاده از فرضیات زیر، توابع انتقال مورد ارزیابی قرار می گیرند:

**فرض اول:** برای شبکه در حالت دولایه ۸ نرون برای لایه مخفی و در حالت سه لایه ۵ نرون در لایه های مخفی در نظر گرفته می شود.

**فرض دوم:** طبق تحقیقات قبلی، بدلیل اینکه الگوریتم آموزش لونیگ-مارکوارت<sup>۱</sup> در امر انتخاب بهترین وزن ها عملکرد بهتری نسبت به سایر الگوریتم ها دارد، انتخاب می شود.

با توجه به دو فرض بیان شده، جدول ۲ بهترین شبکه ها با تغییر توابع انتقال را نشان می دهد.

جدول ۲- مقایسه توابع انتقال مختلف

توابع انتقال	RMSE	R <sup>2</sup>	MAE	MAPE
tansig-purelin	۲۸۹/۲۱	۰/۹۸۵	۱۸۹/۹۵	۵/۷۱۶
logsig-purelin	۲۷۱/۸۵	۰/۹۸۹	۱۹۹/۱۶	۵/۲۷۴
tansig-tansig	۳۲۵/۴۴	۰/۹۸۴	۲۰۸/۶۸	۴/۹۶۵
purelin-purelin	۳۳۳/۷۵	۰/۹۸۴	۲۴۲/۶۳	۵/۳۶۴
tansig-tansig-purelin	۳۵۰/۸۲	۰/۹۸۱	۲۱۸/۲۱	۵/۴۷۶
tansig-logsig-purelin	۳۰۴/۷۳	۰/۹۸۰	۲۱۳/۲	۵/۹۴۷
tansig-purelin-purelin	۲۸۵/۰۲	۰/۹۸۴	۲۰۳/۸۵	۵/۰۸۸
purelin-purelin-purelin	۳۰۷/۹۳	۰/۹۸۴	۲۱۱/۴۲	۴/۹۰۹
logsig-tansig-purelin	۳۱۸/۴۶	۰/۹۸۲	۲۲۰/۷۹	۶/۴۵۹

با توجه به جدول ۲، در شبکه دولایه تابع انتقال لایه مخفی لگاریتم سیگموئید و لایه خروجی تابع انتقال خطی مناسب بوده و در شبکه سه لایه، بترتیب برای لایه های مخفی اول و دوم، توابع انتقال تانژانت سیگموئید و خطی و برای لایه خروجی همانند شبکه دولایه تابع انتقال خطی، بهترین عملکرد را نشان می دهند.

### ۲-۶- ساختار شبکه

<sup>1</sup> Levenberg-Marquardt (LM)



انتخاب بهترین ساختار برای شبکه از اهمیت بالایی برخوردار است، در صورتی که تعداد نرون ها بسیار کم در نظر گرفته شود شبکه خوب آموزش نمی بیند و در صورتی که بسیار زیاد در نظر گرفته شود، امکان بیش برآزش در مدل وجود دارد بنابراین بایستی تعداد نرون ها در مدل را به دقت انتخاب کرد. در نتیجه در این پژوهش، تغییر نرون ها در لایه های مخفی توسط کدنویسی بین ۳ تا ۱۲ نرون انجام می شود و در نهایت با ۱۰ بار اجرا کردن این کد و دریافت RMSE ها بعنوان خروجی، بهترین نتایج RMSE که به صورت کمترین میانگین ها و کمترین واریانس ها ظاهر می شود، در جدول ۳ نشان داده شده است.

همانطور که ملاحظه می شود، بهترین ساختار از نظر کمترین RMSE، متعلق به شبکه با ساختار ۱-۸-۱۰-۶ می باشد. لازم بذکر است توابع انتقال بهینه بدست آمده از بخش ۶-۱ و الگوریتم LM در این بخش مورد استفاده قرار گرفته است.

جدول ۳- مقایسه بهترین ساختار شبکه ها توسط RMSE و واریانس

ساختار شبکه	RMSE	Variance
۱-۱۰-۸-۶	۲۴۳/۳۴	۵۰۷/۴۵۴
۱-۴-۵-۶	۳۱۸/۴۱	۱۶۷/۱۶
۱-۵-۶	۲۵۳/۱۹	۲۲۵/۰۵
۱-۷-۱۱-۶	۲۶۶/۴۷	۳۰۰/۷۱
۱-۵-۱۱-۶	۲۶۸/۶۶	۵۹۸/۸۷
۱-۲-۴-۶	۲۶۸/۵۱	۴۱۲/۶۹
۱-۱۰-۵-۶	۲۹۶/۸۲	۳۳۶/۲۴
۱-۶-۸-۶	۲۶۴/۷۹	۱۰۰۹/۶۲

### ۳-۶- بهترین الگوریتم آموزش

با توجه به نتایج بدست آمده در بخش های ۱-۶ و ۲-۶، بهترین ساختار برای مدل شناسایی شد، حال مهمترین بخش انتخاب بهترین الگوریتم آموزش می باشد. بنابر پژوهش های گذشته در بخش های قبلی برای آموزش شبکه از الگوریتم LM استفاده شد؛ حال در این بخش این الگوریتم با سایر الگوریتم های پس انتشار مهم، مورد مقایسه قرار می گیرد. نتایج این مقایسه در جدول ۴ آمده است.

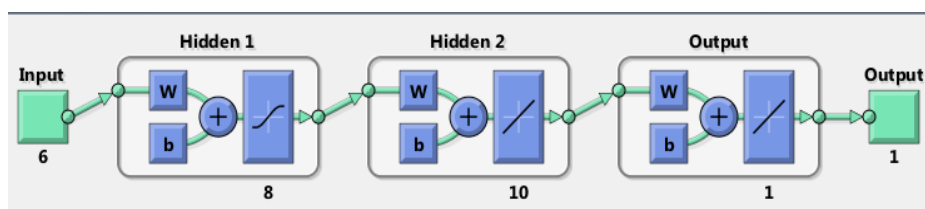
جدول ۴- مقایسه الگوریتم های آموزش

الگوریتم	ساختار شبکه	RMSE	R <sup>2</sup>	MAE	MAPE
lm	۱-۵-۶	۲۵۳/۱۹	۰/۹۸۶	۱۹۶/۱۸	۴/۹۹۱
	۱-۱۰-۸-۶	۲۴۳/۳۴	۰/۹۹	۱۷۷/۹۲	۴/۷۰۶
bfg	۱-۵-۶	۳۷۸/۹۲	۰/۹۷۸	۲۲۴/۳۸	۵/۶۹۵
	۱-۱۰-۸-۶	۳۶۳/۱۸	۰/۹۸	۲۳۰/۶۲	۵/۹۳۱
rp	۱-۵-۶	۳۶۱/۴۳	۰/۹۷۹	۲۴۲/۹۴	۵/۶۶۴
	۱-۱۰-۸-۶	۳۷۹/۳۴	۰/۹۷۹	۲۳۱/۸۹	۶/۵۶

cgf	۱-۵-۶	۳۱۴/۸۱	۰/۹۸۷	۲۳۵/۱۸	۵/۲۱۲
	۱-۱۰-۸-۶	۳۴۵/۹۴	۰/۹۸۳	۲۳۴/۹۱	۶/۳۹۴
scg	۱-۵-۶	۲۹۸/۹۳	۰/۹۸۴	۱۹۳/۶۸	۶/۰۵۵
	۱-۱۰-۸-۶	۳۱۳/۰۱	۰/۹۸۳	۲۵۰/۳۳	۶/۹۳۵
OSS	۱-۵-۶	۳۱۰/۱۶	۰/۹۸۵	۲۳۵/۱	۷/۴۹۴
	۱-۱۰-۸-۶	۳۹۵/۶۶	۰/۹۷۷	۳۰۴/۳۲	۵/۶۳۸

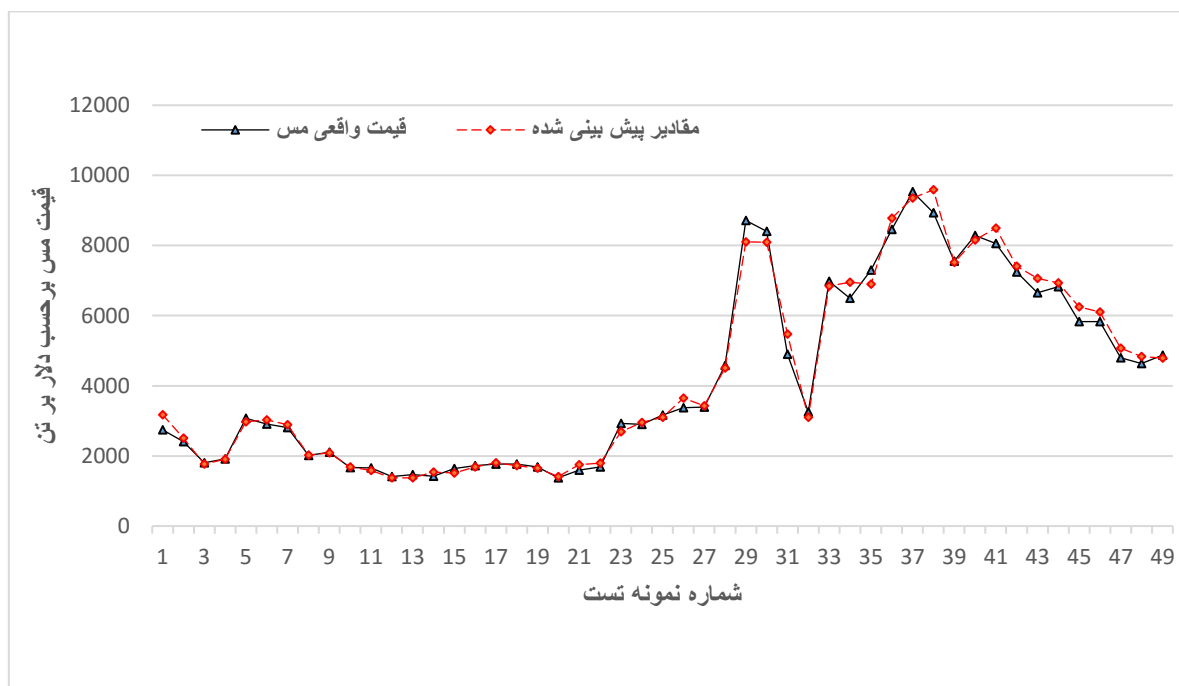
نتیجه بدست آمده در این مقاله نیز برتری الگوریتم LM را نشان می‌دهد.

در نتیجه با توجه به نتایج این بخش بهترین ساختار مدل شبکه عصبی با سه لایه مخفی با ساختار ۱-۱۰-۸-۶ با استفاده از توابع انتقال tansig-purelin-purelin و همچنین الگوریتم آموزش لونیگ-مارکواردت مناسب‌ترین مدل جهت پیش-بینی قیمت مس می‌باشد. شکل ۵ ساختار شبکه پیشنهادی را نشان می‌دهد.



شکل ۵- ساختار مدل پیشنهادی

شکل ۶ نتایج پیش‌بینی قیمت مس توسط داده‌های آموزش را نشان می‌دهد.



شکل ۶-مقایسه نتایج داده‌های آزمون با مقادیر واقعی



## ۷- نتیجه گیری

پیش بینی قیمت مس در میان سرمایه گذاران بازارهای بورس از اهمیت بالایی برخوردار است. علاوه بر انتخاب مدل مناسب جهت پیش بینی، شناسایی و انتخاب پارامترهای موثر بر قیمت مس نیز اهمیت بسیار بالایی در ساخت یک مدل دقیق و کارآمد دارد. از این رو برای افزایش دقت پیش بینی با استفاده از پارامترهای بیشتر و از طرفی اجتناب از تشکیل یک مدل شبکه عصبی پیچیده، روش تحلیل مولفه‌ی اساسی مورد استفاده قرار گرفت. نتایج پیش بینی استفاده از روش PCA جهت ساخت مدل شبکه عصبی را توصیه می نماید. با استفاده از روش ارزیابی RMSE در سه بخش مختلف بهترین ساختار مدل با کمترین RMSE، شبکه عصبی با ساختار ۱-۱۰-۸-۶ می باشد که توابع انتقال در لایه های آن بترتیب شامل tansig، purelin و purelin بوده و همچنین الگوریتم آموزش مناسب آن الگوریتم LM می باشد. RMSE بدست آمده برای این شبکه برابر با ۲۴۳/۳۴ است.





## منابع

- [۱] زراء نژاد، منصور؛ فقه مجیدی، علی؛ رضایی، روح الله؛ ۱۳۸۷. پیش‌بینی نرخ ارز با استفاده از شبکه‌های عصبی مصنوعی و مدل ARIMA. فصلنامه اقتصاد مقداری، دوره ۵، شماره ۴، صفحه ۱۰۷ تا ۱۳۰
- [۲] قاسمی، عبدالرسول. اسدپور، حسن. شاصادقی، مختار. کاربرد شبکه عصبی مصنوعی در پیش‌بینی سریهای زمانی و مقایسه آن با مدل ARIMA. پژوهشنامه بازرگانی، صفحات ۸۷ تا ۱۱۹
- [۳] فتح آبادی، ز، ۱۳۹۰، پیش‌بینی قیمت مس با استفاده از شبکه عصبی و سری زمانی، پایان‌نامه کارشناسی ارشد، دانشگاه تربیت مدرس، تهران.
- [۴] کیا، م، ۱۳۹۵، شبکه های عصبی در MATLAB. چاپ پنجم، تهران، انتشارات دانشگاهی کیان.
- [۵] جوادنژاد، ف، ۱۳۹۰، ارائه مدلی برای پیش‌بینی قیمت نفت خام با استفاده از روش‌های هیبریدی هوش مصنوعی و سری‌های زمانی، پایان‌نامه کارشناسی ارشد، دانشگاه تربیت مدرس، تهران.
- [۶] خدایاری، علی‌اصغر، رحیمی، محمدعلی، ۱۳۸۴، تعیین مدل مناسب برای پیش‌بینی قیمت مس، نشریه دانشکده فنی، جلد ۴۰، شماره ۱، صفحات ۱۳ تا ۲۲، تهران.
- [7] F,S, Lasheras., F,C,Sánchez., A,Krzemień., P,R,Fernández., 2015. "Forecasting the COMEX copper spot price by means of neural networks and ARIMA models". Resources Policy 45, 37-43.
- [8] Smith KL, Bracker K, 2003, "Forecasting Changes in Copper Futures Volatility with GARCH Models Using an Iterated Algorithm", Review of Quantitative Finance and Accounting, 20: 245–265, 2003 Kluwer Academic Publishers. Manufactured in The Netherlands.
- [9] Smith KL, Bracker K, 1999, "Detecting and modeling changing volatility on the copper futures markets, Futures markets 1999, 19, 79-100.
- [10] Bowman C, Husain AM, 2004, " Forecasting Commodity Prices: Futures Versus Judgment", International Monetary Fund.
- [11] Isabel FF, 2007, "Commonality in the LME aluminum and copper volatility processes through a FIGARCH lens. Working Paper 07-02, Business Economics Series 02, Spain.
- [12] Walkins C, McAleer M, 2008, "How has volatility in metals markets changed?" , Mathematics and Computers in Simulation 78 (2008) 237–249.
- [13] Hammoudeh SM, Yuan Y, McAleer M, 2009, " Modeling exchange rate and industrial commodity volatility transmissions" , "MARCO FANNO" WORKING PAPER N.96.
- [14] Chen MH, 2010, " Understanding world metals prices—Returns, volatility and diversification", Resources Policy 35 (2010) 127–140.
- [15] Cortazar G, Eterovic F, 2010, "Can oil prices help estimate commodity futures prices? The cases of copper and silver". Resources Policy 35, 283-291.
- [16] Burcak B, 2010, " Forecasting the prices of non-ferrous metals with GARCH models & volatility spillover from world oil market to non-ferrous metal markets", MBA thesis, Department of Business Administration, 51 pages.
- [17] Javier GC, Roque M, 2010, " Modeling copper price: a regime-switching approach", Pontificia Universidad Católica Argentina.
- [18] Lagos G, 2012, "Directorate for financial and enter price AFFAIRS competition committee DAF/Comp/GF", 1-12.
- [19] Goss BA, Avsar SG, 2013, " Simultaneity, forecasting and profits in London copper futures", University of Adelaide and Flinders University and Wiley Publishing Asia Pty Ltd. 79-96.
- [20] Rutledge RW, Karim K, Wang R, 2013, "International copper futures market price linkage and information transmission: Empirical evidence from the primary world copper markets", International Business Research, Volume 12, Number 1.



[21] Boncic D, Moretto C, 2015, " Forecasting copper prices with dynamic averaging and selection models", North American Journal of Economics and Finance 33 (2015) 1–38.

[22] Bashir ZA, El-Hawary ME, 2009, " Applying Wavelets to Short-Term Load Forecasting Using PSO-Based Neural Networks", IEEE TRANSACTIONS ON POWER SYSTEMS, VOL. 24, NO. 1, 20-27.

[23] Ostadi H, Azimi F, 2015, " Comparing the ability of ANN artificial neural networks and ARIMA method for predicting steel price", Scientific Research and Development 2 (3): 5-10.

[24] Kaastra I, Boyd M, 1996, "Designing a neural network for forecasting financial and economic time series", Neurocomputing 10.

[25] Yari M, Bagherpour R, Jamali S, Shamsi R, 2015, "Development of a novel flyrock distance prediction model using BPNN for providing blasting operation safety", The Natural Computing Applications Forum.

[26] Wikipedia: <https://www.wikipedia.org/> .