Archive of SID

فصلنامه پژوهش‌های اقتصادی ایران / شماره ۱۲/ پاییز ۱۳۸۱

مدل‌سازی و پیش‌بینی رشد اقتصادی در ایران با استفاده از شبکه‌های عصبی مصنوعی(ANN)[1]

محمد رضا قدیمی*
سعید مشیری**

تاریخ ارسال: ۸۱/۷/۱۴ تاریخ پذیرش:۸۱/۱۰/۱۸

چکیده

شبکه‌های عصبی مصنوعی، یک ابزار قدرتمند برای تجزیه و تحلیل داده‌ها و مدل سازی روابط غیر خطی به حساب می آید که استفاده از آنها طی سال‌های گذشته در اقتصاد کلان گسترش یافته است. در این مطالعه، کارایی یک مدل شبکه عصبی با یک مدل خطی رگرسیون برای پیش‌بینی نرخ رشد اقتصادی در ایران مقایسه می شود. برای این منظور ابتدا، یک مدل رگرسیون رشد برای دوره ۱۳۱۵-۱۳۷۳ برآورد شد و سپس با همان مجموعه رگرسورها (متغیرهای ورودی) یک مدل شبکه عصبی طراحی و تخمین زده می شود. پس از آن، با استفاده از معیارهای مرسوم ارزیابی مدل‌های رقیب، کارایی دو مدل فوق در زمینه پیش‌بینی نرخ رشد اقتصادی در ایران در دوره ۱۳۷۴-۱۳۸۰ مقایسه می شوند. نتایج نشان می دهد که مدل شبکه عصبی برای پیش‌بینی نرخ رشد اقتصادی در ایران از کارایی بالاتری برخوردار است.

1. Artificial Neural Networks

* دانشجوی دوره دکترای اقتصاد دانشگاه علامه طباطبائی
** عضو هیئت علمی دانشگاه علامه طباطبائی

Archive of SID
www.SID.ir

Archive of SID

واژه‌های کلیدی : پیش بینی رشد اقتصادی ، شبکه‌های عصبی مصنوعی

Archive of SID

www.SID.ir

Archive of SID

۱. مقدمه

در سال‌های اخیر، پیشرفت‌های قابل توجه در پردازش سریع تر اطلاعات به وسیله کامپیوترها و نرم افزارهای کاربردی، انگیزه پژوهش و به کارگیری مدل‌های غیر خطی را در میان اقتصاددانان به طور چشمگیری افزایش داده است. یکی از معروف ترین مدل‌ها در این زمینه، شبکه‌های عصبی مصنوعی است. مدل‌های شبکه‌های عصبی مصنوعی یکی از ابزارهای قدرتمند تجزیه و تحلیل داده‌ها در علوم مهندسی و کامپیوتر، علوم پزشکی، علوم پایه و بسیاری از رشته‌های مختلف علمی است که استفاده از آن در مدل‌های اقتصاد کلان در دهه ۹۰ مورد توجه پژوهشگران اقتصاد قرار گرفت(البته استفاده از مدل‌های شبکه‌های عصبی مصنوعی در بازارهای مالی سابقه طولانی تری دارد). در واقع مدل‌های شبکه‌های عصبی مصنوعی با استفاده از توابع و پردازشگرهای ریاضی به شبیه سازی عملکرد مغز انسان می پردازند[۱] و قادرند روابط ناشناخته به شدت غیر خطی را مدل‌سازی کنند.

اصلی ترین و مهم ترین کاربرد مدل‌های شبکه‌های عصبی مصنوعی در اقتصاد، پیش‌بینی متغیرهای اقتصادی است[۲]. هدف این مقاله، مقایسه عملکرد یک مدل خطی با مدلشبکه‌های عصبی مصنوعی برای مدل‌سازی و پیش‌بینی رشد اقتصادی در ایران است که از یک مجموعه رگرسورها(به زبان شبکه‌های عصبی متغیرهای ورودی) استفاده می کنند، منظور مدل‌سازی مشخص کردن مدل[۳]، تخمین پارامترها ودر نهایت، ارزیابی مدل برآورد شده است.

---

۱. البته، این جمله کمی اغرارآمیز است زیرا، یک مدل شبکه عصبی مصنوعی به لحاظ پیچیدگی اصلاً قابل مقایسه با مغز انسان نیست. اما به هر حال، علاقه مندی پژوهشگران به راز چگونگی پردازش اطلاعات در مغز انسان و کوشش برای ساختن مدلهای مصنوعی مغز بود که منجر به پدید آمدن شبکه‌های عصبی مصنوعی شد. در هر صورت این مدل‌ها به لحاظ ترکیب (البته در یک الگوی بسیار کوچک تر) شبیه مغز انسان هستند.

۲. دو کاربرد دیگر پذیرفته شده شبکه‌های عصبی مصنوعی در اقتصاد عبارتند از: الف- طبقه بندی واحدهای اقتصادی (Classification of Economic Agents) (مصرف کننده و بنگاه) که به طور عمده، در زمینه پیش‌بینی ورشکستگی واحدهای اقتصادی به کار می رود و ب-مدل‌سازی واحدهای اقتصادی عقلایی محدود شده (Modelling Bounded Rational Economic Agents) که در واقع، فرایند یادگیری واحدهای اقتصادی عقلایی را مدل‌سازی می کند. در این کاربرد، نرونها به عنوان واحدهایی که استنباطشان را از محیط بر اساس اطلاعات دریافتی بهنگام می کنند تفسیر می شوند. سارجنت (Sargent) در سال ۱۹۹۳ برای اولین بار شبکه‌های عصبی مصنوعی را در این حوزه به کار گرفت.

3. Model Specification

www.SID.ir

Archive of SID

در بخش اول این مقاله، مروری بسیار اجمالی بر متدولوژی مدل‌های شبکه‌های عصبی مصنوعی خواهیم داشت. سپس، پژوهش‌های کاربردی در اقتصاد کلان را با استفاده از شبکه‌های عصبی مصنوعی مرور خواهیم کرد. در بخش بعد، کارایی یک مدل خطی رگرسیون را با یک مدل شبکه عصبی مصنوعی برای پیش‌بینی رشد اقتصادی در ایران، مقایسه می کنیم.

۲. مدل شبکه‌های عصبی مصنوعی

یک شبکه عصبی مصنوعی، مجموعه‌ای از نرون‌های به هم متصل در لایه[1]های مختلف است که اطلاعاتی را برای یکدیگر ارسال می کنند. ساده ترین شکل شبکه فقط دو لایه دارد، لایه ورودی[2] و لایه خروجی[3]. شبکه، شبیه یک سیستم ورودی-خروجی عمل می کند و ارزش نرون[4]های ورودی را برای محاسبه ارزش نرون خروجی مورد استفاده قرار می دهد. شکل(۱) نمایش نموداری استاندارد یک شبکه عصبی را نشان می دهد[5]. هر نرون، به وسیله یک دایره و ارتباط میان نرون‌ها با یک فلش نشان داده شده است. خروجی y و ورودی‌ها x_0 ، x_1 و x_2 بردارهای $n\times 1$ هستند که در آن، n تعداد مشاهدات است. در این مثال، اطلاعات منحصراً از ورودی‌ها به خروجی حرکت می کنند، بنابراین، مدل مورد بحث به شبکه عصبی پیش‌خور[6] معروف است[1].

1.Layer
2. Input Layer
3. Outpt Layer

۴. نرون‌ها (Neuron) ساده ترین واحد ساختاری سیستم‌های عصبی طبیعی هستند. در واقع بافت‌هایی که عصب نامیده می شوند، اجتماعی از نرون‌ها هستند. این نرون‌ها، اطلاعات و پیام‌ها را از یک قسمت بدن به قسمت دیگر منتقل می کنند. کلید اصلی درک رفتار مغز انسان به عنوان یک سیستم پردازش اطلاعات، در چگونگی ارتباط واتصال نرون‌ها با یکدیگر است. عنصر اصلی شبکه‌های عصبی مصنوعی نیز، نرون‌های مصنوعی به عنوان عناصر ریاضی پردازشگر هستند. در ادبیات شبکه‌های عصبی، در بسیاری از موارد از کلمه واحد Unit یا گره Node نیز به جای نرون استفاده می شود.

۵. محدودیتی برای تعداد متغیرهای ورودی و خروجی وجود ندارد. در اینجا، برای سادگی فرض شده است سه متغیر ورودی و یک متغیر خروجی وجود دارد.

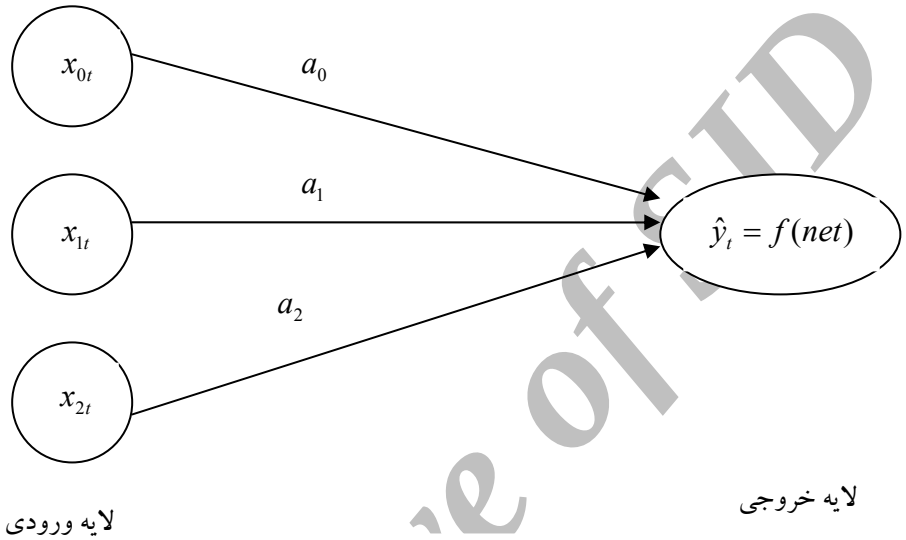
6. Feedforward Neural Network

Archive of SID

www.SID.ir

Archive of SID

شکل-۱. مدل شبکه عصبی ساده پیش‌خور



ارتباط میان یک ورودی و خروجی به وسیله یک وزن a که بیانگر اهمیت نسبی ورودی مذکور در محاسبه ارزش خروجی است، مشخص می شود. به این ترتیب، ارزش نرون خروجی مشاهده t ، از رابطه زیر به دست می آید:

$$Net_t = a_0 x_{0t} + a_1 x_{1t} + a_2 x_{2t} = \sum_{i=0}^{2} a_i x_{it} \qquad (۱)$$

۷. در شبکه‌های پیچیده تر، نرون‌های هر لایه می توانند به یکدیگر مرتبط شوند. همچنین، خروجی‌ها می توانند به ورودی‌ها و یا واحدهای پنهان (در ادامه معرفی می شوند) وصل شوند که طیف گسترده‌ای از مدل‌های همزمان را مشخص می کند.

www.SID.ir

Archive of SID

سپس نرون خروجی، ارزش به دست آمده[۱] را با استفاده از یک *تابع تبدیل یا فعال سازی(محرک)*[۲] که با $f(x)$ نشان داده می شود، پردازش می کند. در ساده ترین شکل شبکه عصبی پیش‌خور، تابع فعال سازی خطی است. برای مثال، $f(x) = x$. ارزش به دست آمده از رابطه (۱) و یک تابع فعال سازی خطی[۳]، خروجی نهایی شبکه برای مشاهده t را به صورت زیر می سازد:

$$\hat{y}_t = f(Net_t = a_0 x_{0t} + a_1 x_{1t} + a_2 x_{2t}) = a_0 x_{0t} + a_1 x_{1t} + a_2 x_{2t} \qquad (۲)$$

معمولاً، یکی از ورودی‌ها برای تمام مشاهدات دارای ارزش یک است و جمله اریب[۴] نامیده می شود. اگر، بپذیریم که x_0 جمله اریب باشد، در آن صورت خروجی شبکه از رابطه زیر به دست می آید:

$$\hat{y}_t = a_0 + a_1 x_{1t} + a_2 x_{2t} \qquad (۳)$$

همان طور که مشاهده می شود، یک شبکه عصبی پیش خور با دو لایه و تابع فعال سازی خطی مشابه *مدل رگرسیون خطی چند متغیره*[۵] است. نرون‌های ورودی همان متغیرهای مستقل یا رگرسورها هستند و نرون خروجی همان برآورد متغیر وابسته است. وزن‌های مختلف شبکه نیز، مشابه پارامترهای مدل رگرسیون و جمله اریب نیز همان عرض از مبدأ یا جمله ثابت در مدل رگرسیون است. در صورتی که وقفه‌های متغیر وابسته را به مجموعه ورودی‌ها اضافه کنیم، در آن صورت به شبکه‌ای مشابه با مدل *اتورگرسیو خطی* AR دست می یابیم.

به طور کلی نقش نرونها در شبکه‌های عصبی، پردازش اطلاعات است و این امر در شبکه‌های عصبی مصنوعی به وسیله یک پردازشگر ریاضی که همان تابع فعال سازی است، انجام می شود. تابع فعال سازی می تواند خطی یا غیر خطی باشد. یک تابع فعال سازی، بر اساس نیاز خاص مسئله‌ای که قرار است به وسیله شبکه عصبی حل شود، از سوی طراح انتخاب می شود. برای مثال، زمانی که

---

۱. ارزش به دست آمده در ادبیات شبکه‌های عصبی، به ورودی خالص(Net Input) معروف است. به همین منظور نیز، از نماد Net برای نشان دادن ارزش خروجی به دست آمده استفاده شده است.

2. Transfer or Activation Function

۱. در ادبیات شبکه‌های عصبی معمولاً منظور از یک تابع فعال سازی خطی استفاده از یک تابع خطی همانی است.

2. Bias

3. Multiple Linear Regression Model

Archive of SID

www.SID.ir

Archive of SID

ارزش‌های خروجی مسئله تنها صفر و یک است، دیگر استفاده از یک تابع فعال سازی خطی مناسب نیست و باید از توابع دیگری که بر اساس مقادیر ورودی مختلف تنها مقادیر صفر و یک را نتیجه می دهند، (مثلاً تابع آستانه ای) استفاده کرد.

نمونه ای که در بالا به آن اشاره شد، برای نرون یا نرون‌های خروجی یک تابع فعال سازی خطی را می پذیرد. برای بهره برداری واقعی از توانایی شبکه‌های عصبی، باید از توابع فعال سازی غیرخطی استفاده کرد. تقریباً تمام شبکه‌های عصبی در بخش‌هایی از شبکه از توابع فعال‌سازی غیر خطی استفاده می‌کنند. این مسئله اجازه می دهد که شبکه الگوهای غیرخطی مناسبی از مجموعه داده‌های پیچیده تولید کند. به صورت ایده آل، تابع فعال سازی باید پیوسته مشتق پذیر و یکنواخت باشد، زیرا، این مسئله عمل پیدا کردن ضرایب مقتضی الگوریتم بهینه یابی را تسهیل می کند. رایج ترین تابع فعال سازی مورد استفاده در ادبیات شبکه‌های عصبی، تابع توزیع تجمعی لجستیک[1] است:

$$f(x) = \frac{1}{1+e^{-x}} \qquad (۴)$$

مقدار تابع لجستیک در محدوده صفر و یک قرار دارد[2]. پژوهشگران شبکه عصبی با استفاده از توابع کراندار، سعی در بازتولید وضعیت فعال سازی نرون واقعی مغز انسان را داشته اند. وقتی که تابع نزدیک به یک است، نرون نسبت به علایم دریافتی بسیار فعال عمل می کند. وقتی تابع نزدیک به صفر است، نرون به ندرت به علایم دریافتی واکنش نشان می دهد.

اگر، متغیری که قصد پیش‌بینی آن را داریم بتواند ارزش‌های منفی را نیز بگیرد، در آن صورت، بهتر است از تابع فعال سازی تانژانت هیپربولیک[3] استفاده کنیم:

$$f(x) = \frac{(e^{x}-e^{-x})}{(e^{x}+e^{-x})} \qquad (۵)$$

تابع تانژانت هیپربولیک، ویژگی‌های یکسانی با تابع لجستیک دارد، اما، مقدار آن در محدوده ۱- و ۱ تغییر می کند.

---

1. Logistic Cumulative Distribution Function

۲. عموماً، تابع فعال سازی دامنه نرون خروجی را محدود می کند و به همین علت، آن را تابع محدودساز (Hard Limitter or Squasher) نیز می نامند.

3. Hyperbolic Tangent

www.SID.ir

Archive of SID

دو تابع فعال سازی معروف دیگر، تابع گوسین[۱] $f(x)=e^{-x^2/2}$ و تابع آستانه‌ای[۲] $f(x)=0$ اگر $x \prec 0$ و $f(x)=1$ در غیر این صورت است.

مجدداً، به شبکه پیش خور ساده شکل (۱) بازمی گردیم. یک تابع فعال سازی لجستیک در نرون خروجی منجر به خروجی زیر برای مشاهده t خواهد شد:

$$\hat{y}_t = f(a_0 + a_1 x_{1t} + a_2 x_{2t}) = \frac{1}{1+e^{-(a_0+a_1x_{1t}+a_2x_{2t})}} \qquad (۶)$$

نتیجه شبکه، مشابه یک *مدل احتمالی لاجیت دوتایی*[۳] است. اگر تابع فعال سازی، یک تابع توزیع تجمعی نرمال باشد، در آن صورت، به یک *مدل پرابیت دوتایی*[۴] دست می یابیم.

وقتی با یک متغیر وابسته که دارای کران نیست کار می کنیم، می توانیم از یک تابع فعال سازی غیرخطی بدون کران مانند $f(x)=x^3$ استفاده کنیم. اما در هر صورت پژوهشگران شبکه‌های عصبی، همان طور که در ادامه به آن پرداخته خواهد شد، ترجیح می دهند از توابع فعال سازی غیرخطی استفاده کنند و با اضافه کردن *لایه‌های پنهان*[۵] به ساختمان[۶] شبکه، به متغیرهای وابسته بدون کران دست یابند.

در کاربردهای واقعی، ساختمان شبکه معمولاً بسیار پیچیده است. همان طور که در شکل(۲) نشان داده شده است پژوهشگران تقریباً همیشه ساختمانی را با یک لایه پنهان،[۷] و یا بیشتر طراحی می کنند. در این شکل، a_{ij} نشانگر وزن ارتباطی است که ورودی i را به نرون یا واحد پنهان j متصل می کند. می‌پذیریم که x_0 یک جمله اریب برای واحدهای پنهان است، درحالی که B یک جمله اریب برای واحد خروجی است.

---

4. Gaussian
5. Threshold
1. Binary Logit Probability Model
2. Binary Probit Model
3. Hidden Layers
4. Architecture

۵. در صورت استفاده از تابع فعال سازی محدود ساز در نرونهای پنهان، این شبکه در ادبیات شبکه‌های عصبی شبکه پرسپترون چند لایه Multilayer Perseptron یا MLP نامیده می شود.

Archive of SID

www.SID.ir

Archive of SID

شبکه‌های عصبی با لا یه‌های پنهان دارای توانایی‌های بیشتری نسبت به شبکه‌های عصبی دو لایه هستند. می توان ثابت کرد که شبکه‌های عصبی پیش‌خور با یک لایه پنهان، تابع فعال سازی لجستیک در لایه پنهان، تابع فعال سازی خطی در نرون خروجی و تعداد نرون‌های کافی در لایه پنهان، قادرند هر تابعی را با دقت دلخواه تقریب بزنند[۱].

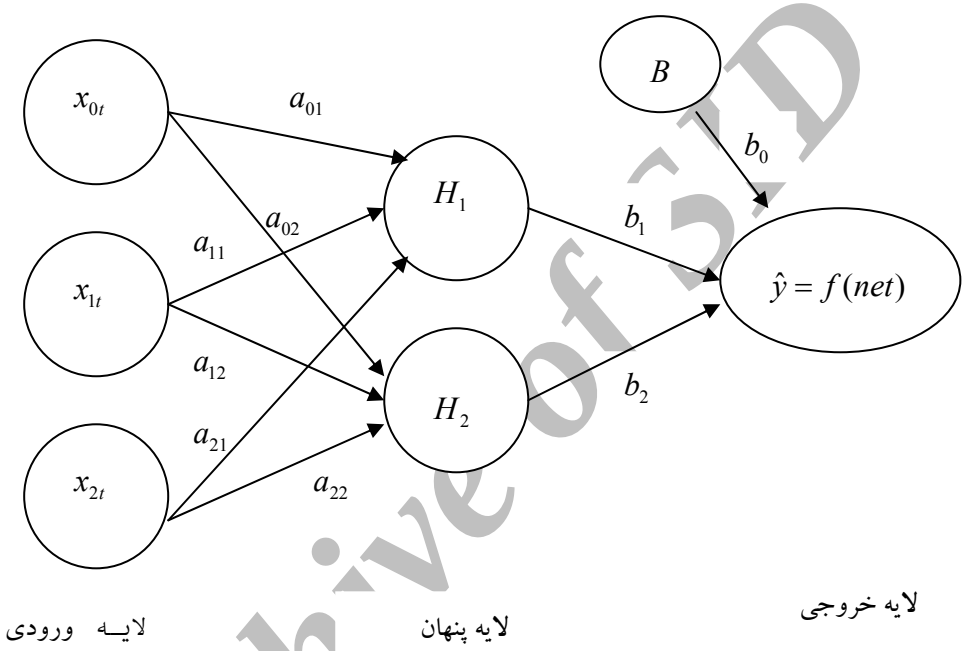
Archive of SID

۶. ابتدا، رامل هارت و دیگران (Rumelhart et al, 1986) و پس از آن، بسیاری از نویسندگان از جمله هچ نیلسون (Hecht-Nielson,1987)، سایبنکو (Cybenko,1989)، فاناهاشی (Funahashi,1989)، هارنیک و دیگران (Hornik et al,1989,1990) و وایت (White,1992) با دقت تمام نشان دادند که یک شبکه عصبی با یک لایه پنهان و تابع فعال سازی لجستیک در واحدهای پنهان، یک تقریب زننده جامع Universal Approximator است. به این معنی که، برای یک تعداد کافی از واحدهای پنهان، شبکه تقریباً می تواند هر تابع خطی یا غیر خطی را با یک سطح دلخواهی از دقت تقریب بزند.

www.SID.ir

Archive of SID

شکل-۲. شبکه عصبی پیش خور با یک لایه پنهان



بر خلاف واحدهای ورودی و خروجی، واحدهای پنهان هیچ مفهومی را نشان نمی دهند. واحدهای پنهان تفسیر یا معنی خاصی ندارند و صرفاً یک نتیجه میانی در فرایند محاسبه ارزش خروجی هستند. بنابراین، آنها هیچ معادلی در ادبیات اقتصاد سنجی ندارند. واحدهای پنهان شبیه واحدهای خروجی رفتار می کنند. برای نمونه، آنها مجموع موزونی از متغیرهای ورودی را محاسبه و سپس، با استفاده از یک تابع فعال سازی که در بیشتر مواقع یک تابع لجستیک است، نتیجه را پردازش

www.SID.ir

Archive of SID

می کنند. در شبکه نمایش داده شده در شکل (۲) ، نتیجه تولید شده به وسیله واحدهای پنهان عبارت است از:

$$H_1 = f(a_{01} + a_{11}x_1 + a_{21}x_2) = \frac{1}{1+e^{-(a_{01}+a_{11}x_1+a_{21}x_2)}}$$
$$H_2 = f(a_{02} + a_{12}x_1 + a_{22}x_2) = \frac{1}{1+e^{-(a_{02}+a_{12}x_1+a_{22}x_2)}} \quad (۷)$$

با جایگزینی تابع فعال سازی لجستیک درواحدهای پنهان به جای واحد خروجی، شبکه دیگر به تولید تخمین متغیرهای کراندار محدود نباشد. اگر متغیر وابسته دارای کران نیست، واحد خروجی معمولاً از یک تابع فعال سازی خطی استفاده می کند. برای نمونه، خروجی شبکه برابر مجموع موزون ارزشهای واحدهای پنهان با ضرایب b_j خواهد بود. به این ترتیب، یک خروجی پیوسته، غیرخطی و بدون کران به شکل زیر حاصل خواهد شد:

$$\hat{y} = b_0 + b_1H_1 + b_2H_2$$
$$\hat{y} = b_0 + \frac{b_1}{1+e^{-(a_{01}+a_{11}x_1+a_{21}x_2)}} + \frac{b_2}{1+e^{-(a_{02}+a_{12}x_1+a_{22}x_2)}} \quad (۸)$$

اگر متغیر وابسته کراندار باشد، واحد خروجی نیز معمولاً برای پردازش از یک تابع فعال سازی لجستیک استفاده می کند، بنابراین، یک خروجی کراندار همانند معادله (۹) به دست خواهد آمد:

$$\hat{y} = f(b_0 + b_1H_1 + b_2H_2)$$
$$\hat{y} = \frac{1}{1+e^{-(b_0+b_1H_1+b_2H_2)}} \quad (۹)$$
$$\hat{y} = \frac{1}{1+e^{-\left(b_0+\frac{b_1}{1+e^{-(a_{01}+a_{11}x+a_{21}x_2)}}+\frac{b_2}{1+e^{-(a_{02}+a_{12}x+a_{22}x_2)1}}\right)}}$$

Archive of SID

www.SID.ir

Archive of SID

بدین ترتیب، شبکه ضمن حفظ ویژگی‌های غیرخطی مدل، می تواند خروجی‌های کراندار و یا بدون کران تولید کند که نتیجه آن شبیه *رگرسیون غیرخطی چندمتغیره* است.

اگر ارتباط مستقیمی از ورودی‌ها به خروجی برقرار شود، به مدل جدید دیگری می توان دست یافت. در این ساختمان که به شبکه عصبی بهبود یافته[1] یا شبکه عصبی با ارتباطات پرشی[2] معروف است، ورودی‌ها به طور مستقیم به خروجی وصل می شوند.

اگر تابع فعال سازی خطی را برای نرون خروجی بپذیریم، شبکه عصبی بهبود یافته دارای ساختمان بسیار جالبی می شود، زیرا، مدل رگرسیون خطی را نیز در بر می گیرد. بنابراین، رایج ترین ساختمان برای پیش‌بینی متغیرهای اقتصاد کلان با استفاده از شبکه‌های عصبی است. با استفاده از معادله(۸)، می‌توان خروجی شبکه بهبود یافته را به صورت مستقیم استخراج کرد:

$$\hat{y} = b_0 + a_{1\hat{y}}x_1 + a_{2\hat{y}}x_2 + \frac{b_1}{1+e^{-(a_{01}+a_{11}x_1+a_{21}x_2)}} + \frac{b_2}{1+e^{-(a_{02}+a_{12}x_1+a_{22}x_2)}} \qquad (۱۰)$$

بنابراین، شبکه عصبی بهبود یافته را می توان به عنوان مدل رگرسیون خطی استاندارد که با عبارات غیرخطی توسعه داده شده است در نظر گرفت. اگر متغیر وابسته مورد نظر هیچ نوع ویژگی غیرخطی از خود نشان ندهد، ضرایب b_1 و b_2 برابر صفر خواهد بود و به یک مدل خطی استاندارد خواهیم رسید.

۳. روش‌های تخمین شبکه‌های عصبی مصنوعی

در ادبیات شبکه‌های عصبی، به جای اصطلاح تخمین ضرایب از اصطلاح *یادگیری*[3] یا *آموزش*[4] برای پیدا کردن ارزش‌های وزن‌های شبکه استفاده می شود. دو نوع یادگیری در این ادبیات مورد بحث قرار

1. Augmented Neural Networks
1. Jump Connections
2. Learning
3. Training

Archive of SID

www.SID.ir

Archive of SID

می‌گیرد: *یادگیری تحت نظارت*[۱] و *یادگیری بدون نظارت*[۲]. یادگیری با نظارت که به یادگیری با معلم نیز معروف است، ارزش‌های متغیر هدف که شبکه باید بر اساس ارزش‌های متغیرهای ورودی از طریق محاسباتش آنها را دوباره تولید کند، مشخص و سپس، خطای پیش‌بینی برای هر مشاهده، به وسیله محاسبه اختلاف خروجی شبکه با ارزش‌های متغیرهای هدف اندازه گیری می شود. پس از آن، با استفاده از الگوریتم‌های مختلف *تکرار*[۳] (که مشهورترین آنها الگوریتم *پس انتشارخطا*[۴] است) وزن‌های شبکه تعدیل می شود(اصطلاحاً شبکه آموزش داده می شود)، به گونه‌ای که خطای پیش‌بینی داخل نمونه که به وسیله مجموع مربعات خطاها یا میانگین خطای مطلق اندازه گیری می شود حداقل شود. همین طور که وزن‌ها با هر تکرار تغییر می کند(شبکه آموزش داده می شود) اصطلاحاً گفته می شود که شبکه در حال *یادگیری* است (در این مقاله، هماهنگ با ادبیات اقتصاد سنجی غالباً از اصطلاح تخمین ضرایب به جای یادگیری برای تعیین وزن‌های شبکه استفاده خواهد شد).

تخمین ضرایب یک شبکه عصبی به عنوان یک سیستم به شدت غیر خطی، به آسانی تخمین پارامترهای مدل‌های خطی نیست. ممکن است جواب‌های بهینه نسبی متعددی برای حداقل کردن اختلاف بین ارزش حقیقی متغیر خروجی و ارزش‌های به دست آمده از شبکه وجود داشته باشد که هیچ کدام از آنها بهترین جواب نباشند.

مجموعه اولیه از وزن‌ها می تواند در هر جایی روی محور x ، نزدیک به یک مینیمم نسبی، ماکزیمم مطلق یا نسبی و یا یک نقطه زینی، به جای یک مینیمم مطلق قرار بگیرد. همان طور که وزن‌ها تعدیل می‌شوند، ممکن است در هر کدام از این موقعیت‌ها که در آن مشتق برابر صفر است به دام بیفتیم. بنابراین، تخمین شبکه‌های عصبی زمان بیشتری می برد و استفاده از روش‌های گوناگون را طلب می کند.

برای به دست آوردن ضرایب بهینه شبکه، باید تابع هدف $E(w) = \sum (y - \hat{y})^2$ که در واقع مجموع مربع خطاها یعنی اختلافات میان ارزش‌های خروجی هدف و خروجی به دست آمده از شبکه

4. Supervised Learning
5. Unsupervised Learning
6. Itration
7. Error Backpropagation

Archive of SID

www.SID.ir

Archive of SID

را نشان می دهد، حداقل شود. در الگوریتم‌های تکرار، کار با یک حدس اولیه درباره ضرایب w آغاز می‌شود. در این حالت، رابطه وضعیت مجموعه بهینه ضرایب یا وزن‌ها از شکل زیر پیروی می کند:

$$w^{t+1} = w^t + \Delta w^t \qquad (۱۱)$$

که در آن، t رتبه تکرار و w^t تخمین فعلی از نقطه بهینه، یعنی مینیمم تابع $E(w)$ در مرحله tام است. رابطه بالا را به صورت زیر نیز می توان نوشت:

$$\Delta w^t = w^{t+1} - w^t = \eta p^t \qquad (۱۲)$$

که در آن، p^t بردار جستجو است. به این معنی که از روی آن، تخمین مرحله بعد w^{t+1} به دست می‌آید. η نیز، یک مقدار اسکالر بزرگتر از صفر است که نرخ یادگیری[1] (نرخ تنظیم) نامیده می شود و طول قدم در هر تکرار را در مسیر p^t تعیین می کند.

چگونگی انتخاب p^t، روش‌های مختلف مینیمم سازی را از هم متمایز می کند. بردار جستجو را می‌توان از روی اطلاعات به دست آمده از بردار گرادیان(روش پس انتشار خطا BP) و ماتریس هشین تابع خطا $E(w)$(روش نیوتن[1]، روش شبه نیوتن[2] و روش لونبرگ-مارکوانت[3] LM) تعیین کرد[4].

بعد از تخمین مدل، برای ارزیابی اجرای مدل‌های رقیب، باید قدرت پیش‌بینی مدل‌ها در خارج از نمونه[5] را بررسی کرد. برای این منظور معمولاً، داده‌ها را به دو مجموعه جدا تقسیم می کنند. بخش اول به مجموعه آموزش[6] یا تخمین و مجموعه دوم به مجموعه آزمون[7] موسوم است.

---

1. Learning Rate
1. Newton Method
2. Quasi-Newton Method
3. Lunberg Marquant

۴. روش‌های مذکور به روش‌های گرادیان نزولی (Gradient-Descent Methods) معروفند. در کاربردهای جدیدتر از روش دیگری بنام الگوریتم ژنتیک (Genetic Algorithm) نیز استفاده می شود. روش الگوریتم‌های ژنتیک یکی دیگر از روش‌های برنامه ریزی است که ایده‌های اصلی خود را از زیست شناسی قرض گرفته است. ایده ساده این الگوریتم‌ها، تقلید از مسیر تکامل در طبیعت، برای یافتن یک جواب برای یک مسئله مشخص است. این روش به ویژه در وضعیت‌هایی که در آن فضای جستجوی جواب‌های ممکن ابعاد بسیار بزرگی (High-Dimensional)، دارد مفید است.

5. Out of Sample
6. Training Set

Archive of SID

www.SID.ir

Archive of SID

ابتدا ،ضرایب مدل با استفاده از داده‌های مجموعه اول تخمین زده می شود و سپس، با استفاده از داده‌های مجموعه دوم قدرت پیش‌بینی مدل و یا به عبارت دیگر، توان تعمیم مدل به خارج از مجموعه داده‌های مورد استفاده در تخمین ارزیابی می شود. در این قسمت هدف حداقل کردن خطای پیش‌بینی در مجموعه آزمون است. برای این منظور معمولاً از چهار معیار زیر استفاده می شود[۲]:

الف) معیار میانگین مربع خطا MSE یا ریشه میانگین مربع خطا $RMSE$[۳]:

$$MSE = \frac{\sum(\hat{y}_t - y_t)^2}{n} \quad \text{و} \quad RMSE = \sqrt{\frac{\sum(\hat{y}_t - y_t)^2}{n}} \qquad (۱۳)$$

ب) معیار میانگین قدرمطلق انحراف[۴] (MAD) یا میانگین قدر مطلق درصد خطا[۵] ($MAPE$):

$$MAD = \frac{\sum|\hat{y}_t - y_t|}{n} \quad \text{و} \quad MAPE = \frac{\sum\left|\frac{\hat{y}_t - y_t}{y_t}\right|}{n} \qquad (۱۴)$$

ج) آماره U تایل[۶]:

آماره U تایل، یک معیار اندازه‌گیری بر اساس $RMSE$ است. آماره U، به صورت نسبت ریشه میانگین مجذور خطای مدل پیش‌بینی به ریشه میانگین مجذور خطای یک مدل ابتدایی[۱] که می‌پذیرد

---

7. Test Set

۸. سه معیار اول در هلدن، پیل و تامپسون(Holden,Peel & Thompson,1990) و معیار چهارم در سوآنسون و وایت (Swanson & White,1997) مورد بحث قرار گرفته است.

9. Mean Squared Error or Root Mean Squared Error

1. Mean Absolute Deviation

2. Mean Absolute Persentage Error

3. Theil U Statistic

Archive of SID

www.SID.ir

Archive of SID

ارزش متغیر مورد پیش‌بینی تغییر نخواهد کرد، محاسبه می‌شود. بنابراین، پیش‌بینی در دوره بعد به صورت ساده، ارزش واقعی دوره قبل را اختیار می کند.

$$U = \frac{\sqrt{\frac{\sum(\hat{y}_t - y_t)^2}{n}}}{\sqrt{\frac{\sum(y_{t-1} - y_t)^2}{n}}} \qquad (۱۵)$$

هراندازه آماره U به صفر نزدیک تر باشد، بهتر است. معمولاً، ارزش‌های کوچکتر از 0.55 مورد قبول واقع می شوند.مدل‌هایی که آماره U آنها بزرگ‌تر از یک باشد رد می شوند، زیرا در این حالت پیش‌بینی مدل از یک مدل ابتدایی نیز بدتر بوده است.

د) نرخ اغتشاش[۲]: این معیار، درستی جهت پیش‌بینی‌ها را به وسیله محاسبه نسبت تعداد ناصحیح جهت پیش‌بینی‌ها به تعداد کل پیش‌بینی‌ها اندازه می گیرد. مقدار این معیار، بین صفر و یک است. اگر تمام جهت پیش‌بینی‌ها صحیح باشد، انداز معیار برابر صفر و اگر همه آنها غلط باشد، برابر یک خواهد شد.
در پایان این بخش، به این نکته اشاره می کنیم که تجربه نشان داده است، شبکه‌های عصبی علی رغم اینکه ممکن است در محدوده داده‌های مورد استفاده برای تخمین، نتایج خوبی را به نمایش بگذارند(خطای کمی داشته باشند)، اما، بسیار مستعدند که هنگام استفاده از داده‌های جدید، با خطای زیادی روبه رو شوند و لذا ، نتایج ضعیفی برای پیش‌بینی‌های خارج از نمونه ارایه کنند[۳]. برای حداقل

4. Naïve Model
5. Confusion Rate

۱. این مشکل به عنوان مشکل overfitting شناخته می شود.

Archive of SID

www.SID.ir

Archive of SID

کردن این معضل چندین روش ارایه شده است که رایج‌ترین آنها روش توقف زود هنگام[۱] و منظم سازی بیزین[۲] است[۳].

۴. مروری بر کاربردهای تجربی مدل‌های شبکه‌های عصبی مصنوعی در اقتصاد

موفقیت کم نظیر شبکه‌های عصبی به عنوان ابزاری قدرتمند به منظور تجزیه و تحلیل داده‌ها در علوم تجربی، موجب شد تا توجه اقتصاد دانان نیز به این روش مدل‌سازی جلب شود. ابتدا، با توجه به نیاز این مدل‌ها به داده‌های زیاد، پژوهش و مدل‌سازی در حوزه اقتصاد با استفاده از شبکه‌های عصبی در بازارهای مالی آغاز شد و در اواخر دهه ۸۰ میلادی مدل‌های مختلفی به منظور پیش‌بینی نرخ‌های ارز، قیمت سهام و شاخص‌های مختلف بورس ساخته شد[۴]. از جمله این کارها می توان به پژوهش وایت (۱۹۸۸) که یک شبکه عصبی سه لایه را بر روی ۱۰۰۰ داده از قیمت سهام شرکت *IBM* به کار گرفت اشاره کرد. هدف وایت به جای پیش‌بینی، آزمون فرضیه کارایی بازار بود. او، نتوانست شواهدی را در مقابل نظریه‌ای که پیشنهاد می کند یک گام تصادفی[۵] بهترین مدل برای پیش‌بینی بازارهای مالی است بیابد. اما، شبکه استفاده شده وایت بسیار ساده بود و همین امر موجب شد نویسندگان زیادی نتایج وی را به مجادله بگیرند و با استفاده از شبکه‌های پیچیده تر نشان دهند که فرایند غیر خطی معنی داری در بسیاری از سری‌های زمانی مالی وجود دارد. از جمله این کارها می توان به پژوهش‌های وانگ[۶] (۱۹۹۰)، تریپی و توربان[۷] (۱۹۹۰)، بوزارج[۸] (۱۹۹۳)، تسیبوریس و زیدنبرگ[۱] (۱۹۹۵)،ریفنس[۲] (۱۹۹۵)، ریفنس و دیگران(۱۹۹۵)، هیمسترا[۳] (۱۹۹۶) و هافک و هلمنستین[۴] (۱۹۹۶) اشاره کرد.

---

2. Early Stopping
3. Bayesian Regularization

۴. سارل (Sarle, 1995) روشهای معمول دیگر حداقل کردن مشکل overfitting را مورد بحث قرار داده است. ریچ (Rech,2002) نیز یک طبقه بندی مناسب از این روشها را ارایه کرده است.

۵. همان طوری که قبلاً عنوان شد، شبکه‌های عصبی قادرند هر تابع غیر خطی پیوسته أی را در صورتی که در لایه پنهان به تعداد کافی واحدهای پردازش اطلاعات وجود داشته باشد، تقریب بزنند. بنابراین آنها می توانند مدلهای غیر خطی کارایی را برای سریهای زمانی مالی فراهم کنند و پیش‌بینی‌های بهتری ارایه دهند.

6. Random Walk
1. Wong
2. Trippi & Turban
3. Bosarge

www.SID.ir

Archive of SID

موفقیت شبکه‌های عصبی در حوزه اقتصاد مالی، توجه متخصصان اقتصادکلان و اقتصادسنجی را نیز به خود جلب کرد و پژوهش در زمینه استفاده از شبکه‌های عصبی برای پیش‌بینی و مدل‌سازی در اقتصاد کلان در دهه ۹۰ آغاز شد. اوج این پژوهشها را می توان به دوره پس از انتشار مقاله مشهور خوان و وایت[5] (۱۹۹۴) نسبت داد. خوان و وایت در مقاله خود بسیاری از موضوعات موازی در شبکه‌های عصبی و اقتصاد سنجی (که به بخش‌هایی از آن در بخش قبل اشاره شد) را استخراج کردند و از این دیدگاه، مقاله آنها یک مقدمه پذیرفته شده و قطعی در ادبیات اقتصاد سنجی به حساب می‌آید. پیشبرد نظری خوان و وایت در سال ۱۹۹۴ با انجام یک پژوهش کاربردی در همان سال به وسیله معصومی، ختن زاد و ابایی[6] دنبال شد. این نویسندگان نشان دادند ۱۴ سری زمانی اقتصاد کلان که نلسون و پلاسر[7] در مقاله اثرگذار خود در سال ۱۹۸۲ مورد تجزیه وتحلیل قرار دادند، به خوبی با استفاده از شبکه‌های عصبی قابل مدل‌سازی هستند، بدون این شک قوی که این سری‌ها از فرایند ریشه واحد پیروی می کنند. در حقیقت، این موضوع با نتایج پرون[8] (۱۹۸۹) که نشان داد یک یا دو تحول ساختاری موجب ایستایی چنین سری‌هایی می شوند، سازگار است. در واقع، قدرت اصلی شبکه‌های عصبی این است که آنها قادرند برای تحولاتی چون سقوط بازار بورس و شوک‌های نفتی به عنوان انحرافات معنی دار از قبول فرض خطی بودن، مدل‌های بهتری بسازند.

هیل[9] و دیگران (۱۹۹۴) مجموعه‌ای از مقاله‌های تجربی وکاربردی را برای مقایسه نتایج پیش‌بینی شبکه‌های عصبی و مدل‌های آماری مورد بررسی قرار دادند. در مطالعات مورد بررسی آنها، شبکه‌های عصبی در زمینه پیش‌بینی متغیرهای اقتصاد کلان، با توجه به درصد میانگین قدرمطلق خطا به خوبی مدلهای آماری استاندارد و یا بهتر از آنها عمل کردند. در کاربردهای سری‌های زمانی، نتایج

4. Tsibouris & Zeidenberg
5. Refenes
6. Hiemstra
7. Haefke & Helmenstein
8. Kuan & White
9. Maasoumi, Khotanzad & Abaye
10. Nelson & Plosser
11. Perron
12. Hill

Archive of SID

www.SID.ir

Archive of SID

مقاله‌ها پیشنهاد می کرد که شبکه‌های عصبی برای پیش‌بینی با افق زمانی طولانی تر نتایج دقیق‌تری را به همراه دارد. آنها همچنین، با داده‌هایی با تکرر بیشتر(داده‌های ماهانه یا فصلی) نتایج بهتری را ارایه می کنند. این موضوع، نویسندگان را به این تفکر واداشت که داده‌هایی با تکرر بیشتر دارای فرایندهای غیرخطی بیشتری هستند[1]. در هر صورت، نویسندگان نتیجه گرفتند که مقاله‌های نگارش شده در زمینه موضوع هنوز غیر قطعی هستند.

مقاله کهزادی[2] و دیگران(۱۹۹۵)، یک مقدمه خوب دیگر برای شبکه‌های عصبی و کاربردهای آنها در اقتصاد است. این نویسندگان ضمن پیش‌بینی معاملات سلف ذرت، از مقایسه یک شبکه عصبی با یک مدل *ARIMA* ، با استفاده از معیارهای مختلف عملکرد پیش‌بینی دریافتند که خطای پیش‌بینی مدل شبکه عصبی بین ۱۸ تا ۴۰ درصد کمتر از مدل *ARIMA* است.

این چشم انداز نسبی مثبت برای شبکه‌های عصبی، به وسیله سه مطالعه دیگر که به پیش‌بینی تولید اقتصادی پرداخته اند تأیید شده است. کاکز[3] (۱۹۹۹) صحت مدل‌های خطی و شبکه‌های عصبی برای پیش‌بینی رشد *GDP* واقعی کانادا را با استفاده از سری‌های شاخص‌های مالی مقایسه کرد. در افق زمانی یک فصل بعد و یک سال بعد، شبکه‌های عصبی پیش‌بینی‌های خارج از نمونه دقیق‌تری از مدل‌های خطی ارایه کردند. بر اساس آزمون‌های مختلف، بهبود در دقت پیش‌بینی به دست آمده به وسیله شبکه‌های عصبی به طور کلی از نظر آماری معنی دار بودند. نویسنده نتیجه می گیرد که شبکه‌ها می توانند فرایندهای غیر خطی در ارتباط میان رشدتولید ناخالص داخلی واقعی و شاخص‌های مالی را نشان دهند.

به طور مشابه فو[4] (۱۹۹۸) دریافت که برای پیش‌بینی‌های خارج از نمونه رشد تولید ناخالص داخلی واقعی ایالات متحده، شبکه‌های عصبی بهتر از مدل‌های رگرسیون خطی عمل می کنند. در این مطالعه، شبکه‌های عصبی مجموع مربع پسماندهای خارج از نمونه را بین ۱۰ تا ۲۰ درصد کاهش دادند.

---

۱. البته، هیچ دلیلی وجود ندارد که بپذیریم همه سری‌های زمانی با تکرر بیشتر لزوماً فرایندهای غیرخطی بیشتری را در بر می گیرند. اما در هر صورت، در تکرارهای بیشتر، بعضی از سری‌های مالی یک توزیع شرطی غیر گوسین دارند و بعضی دیگر همچنین یک رفتار متقارنی را از خود نشان می دهند که بوسیله شبکه‌های عصبی بهتر قابل مدل‌سازی هستند. این موضوع، به عنوان یک مسئله تجربی باقی می ماند.

2. Kohzadi
3. Tkacz
1. Fu

Archive of SID

www.SID.ir

Archive of SID

مودی، لوین و رفوس[1] (۱۹۹۳) نیز برای پیش‌بینی نرخ رشد شاخص تولید صنعتی ایالات متحده نتایج مشابهی به دست آوردند. در این پژوهش، برای تمام افق‌های پیش‌بینی مورد بررسی(یک تا دوازده ماه بعد) شبکه‌های عصبی آنها، پیش‌بینی‌های دقیق‌تری از مدل اتورگرسیو و مدل رگرسیون خطی ارایه کردند.

یک مطالعه دیگر در زمینه پیش‌بینی نرخ تورم کانادا نیز، نتایج مفیدی را برای استفاده از شبکه‌های عصبی به همراه داشت. مشیری و کامرون[2] (۲۰۰۰) عملکرد شبکه‌های عصبی مصنوعی با سایر روش‌های اقتصاد سنجی سنتی و سری‌های زمانی را برای پیش‌بینی نرخ تورم کانادا مقایسه کردند. در این مقاله، آنها مدل‌های شبکه‌های عصبی را با مدل‌های ساختاری، VAR ، $BVAR$ و $ARIMA$ برای افق‌های زمانی مختلف(یک، سه و دوازده ماه بعد) مقایسه می کنند. معیارهای به کار گرفته شده در این پژوهش، ریشه میانگین مربع خطاها ($RMSE$) و میانگین قدرمطلق خطاها (MAE) نشان دادند که شبکه‌های عصبی قادرند به خوبی تمام روش‌های اقتصاد سنجی سنتی و سری‌های زمانی، و در بعضی از موارد بهتر از آنها نرخ تورم کانادا را پیش‌بینی کنند.

در مطالعه دیگری درباره تورم کانادا، مشیری و دیگران (۱۹۹۹)، ضمن مقایسه نتایج شبکه‌های عصبی با مدل‌های سری زمانی و اقتصاد سنجی، عملکرد شبکه‌های عصبی مرسوم پس انتشار خطا (منظور شبکه‌هایی است که از الگوریتم پس انتشار خطا برای تخمین استفاده می کنند)، را با شبکه‌های عصبی بازگشتی و شبکه‌های RBF [3] مقایسه می کنند[4]. در این پژوهش نیز آنها برای مقایسه نتایج از دو معیار ریشه میانگین مربع خطاها ($RMSE$) و میانگین قدرمطلق خطاها (MAE) استفاده‌می‌کنند و مدل‌های مختلف شبکه‌های عصبی را با یک مدل اقتصاد سنجی ساختاری

---

2. Moody, Levin & Rehfuss
3. Moshiri & Cameron
4. Radial Basis Function

۵. در شبکه‌های عصبی بازگشتی امکان بازخورد وجود دارد و واحدهای خروجی و پنهان می توانند به ارزش‌های باوقفه خودشان وابسته باشند. شبکه‌های RBF نیز از ساختاری مشابه شبکه‌های عصبی معمول برخوردارند، اما، فرایند پردازش اطلاعات و الگوریتم یادگیری در واحدهای پنهان اساساً با شبکه‌های عصبی معمول متفاوت است. خوانندگان علاقه مند می توانند یک مرور روان بر جزییات چگونگی عملکرد شبکه‌های RBF را در مقاله مشیری و دیگران با نشانی گفته شده در فهرست منابع بیابند.

Archive of SID

www.SID.ir

Archive of SID

یک مدل VAR برای پیش‌بینی نرخ تورم کانادا در سه افق زمانی مقایسه می کنند. مطالعه مشیری و دیگران نشان داد که در پیش‌بینی‌های ایستا[1] مدل‌های شبکه‌های عصبی مرسوم(پس انتشار خطا) نتایجی هم سطح و یا بهتر از دیگر شبکه‌ها و مدل‌های اقتصاد سنجی دارند. اما، برای پیش‌بینی‌های پویا، شبکه‌های عصبی پس انتشار خطا به طور کلی عملکردی بهتر از مدل اقتصاد سنجی ساختاری ارایه کردند، اما، نسبت به مدل VAR ، به جز برای افق زمانی سه ماه بعد عملکرد ضعیف‌تری داشتند. در مقایسه با مدل شبکه بازگشتی نیز، شبکه پس انتشار خطا فقط برای افق زمانی یک ماه بعد نتایج بهتری را ارایه می کند. برای افق‌های زمانی سه ماه و ۱۲ ماه بعد شبکه‌های بازگشتی پیش‌بینی‌های دقیق‌تری تولید می کنند. مقایسه شبکه‌های عصبی پس انتشار خطا با شبکه‌های RBF نیز نشان داد که برای افق‌های زمانی کوتاه (یک و سه ماه بعد) عملکرد شبکه‌های پس انتشار خطا بهتر است، اما برای افق زمانی یک سال بعد شبکه‌های RBF نتایج بهتری به همراه دارند.

در حوزه بازارهای مالی نیز، چندین مطالعه با نتایج مطلوب با استفاده از شبکه‌های عصبی انجام شده است. در اینجا، تنها دو مطالعه که روی متغیرهای مورد علاقه اقتصاد کلان دانان متمرکز شده ارایه خواهد شد. ورکویجن[2] (۱۹۹۶)، دقت مدل‌های مختلف برای پیش‌بینی نرخ مبادله دلار آمریکا به مارک آلمان در افق‌های زمانی ۱ تا ۳۶ ماه بعد را با یکدیگر مقایسه می کند. برای پیش‌بینی‌های خارج از نمونه، مدل‌های شبکه‌های عصبی اندکی دقیق‌تر از مدل‌های رگرسیون خطی و پیش‌بینی‌های گام تصادفی، به ویژه در افق‌های زمانی طولانی تر بودند. عملکرد شبکه‌های عصبی حتی برای پیش‌بینی جهت تغییرات نرخ ارز نیز از سایر مدل‌ها بهتر بود.

در مطالعه دیگر، رفنس، زاپرانیس و فرانسیس[3] (۱۹۹۴) دقت یک شبکه عصبی پیش‌خور و یک مدل رگرسیون خطی را برای پیش‌بینی عملکرد سهام در چارچوب نظریه قیمت گذاری آربیتراژ مقایسه می کنند. نتایج نشان دادند که برای پیش‌بینی‌های داخل و خارج از نمونه، شبکه‌های عصبی دقیق‌ترند[1].

---

۱. در پیش‌بینی ایستا، از ارزش‌های واقعی تاریخی متغیرها برای پیش‌بینی دوره‌های بعد استفاده می شود، در حالی‌که در پیش‌بینی پویا از ارزش‌های پیش‌بینی‌های قبلی برای پیش‌بینی دوره‌های بعد استفاده می شود و خطای پیش‌بینی تولید شده در یک دوره، پیش‌بینی و خطای پیش‌بینی تولید شده برای همه دوره‌های بعد در افق زمانی را تحت تأثیر قرار می دهد.

2. Verkooijen
1. Refenes, Zapranis & Francis

www.SID.ir

Archive of SID

۵. مدل‌سازی و پیش‌بینی رشد اقتصادی در ایران

مطالعه جنبه‌های مختلف رشد اقتصادی در ایران به دو دلیل عمده با اهمیت است. نخست آنکه، سیاست‌گزاران و تصمیم گیران کشور برای اتخاذ تصمیمات اقتصادی مناسب ضروری است که نسبت به چگونگی شکل گیری پدیده رشد اقتصادی در ایران و عوامل مؤثر بر آن شناخت کافی داشته باشند. از این رو، هرگونه مطالعه ای که در یک چارچوب علمی به تبیین این موضوع بپردازد، در نوع خود مفید و مهم به شمار می آید. دوم آنکه، کارگزاران و اقتصاد گردانان کشور- اعم از دولتی و خصوصی- برای برنامه ریزی در سطح کشور و یا حتی در سطح یک بنگاه اقتصادی نیاز دارند که با مدل‌سازی‌های مناسب نرخ رشد اقتصادی را برای مقاطع مختلف زمانی پیش‌بینی کنند. هر چند که این دو دسته مطالعه به شدت به یکدیگر وابسته هستند اما، باید تأکید کرد که هدف اولیه و تمرکز ما در مطالعه حاضر، مدل‌سازی و پیش‌بینی رشد اقتصادی در ایران است.

برای این منظور ابتدا، یک مدل خطی رگرسیون رشد و سپس، یک مدل شبکه عصبی را تخمین می زنیم و کارایی دو مدل را برای پیش‌بینی نرخ رشد اقتصادی در ایران با استفاده از معیارهای مرسوم مورد ارزیابی قرار می دهیم.

۵-۱. داده‌های مورد استفاده

با توجه به عدم وجود داده‌های فصلی رسمی، و تفاوت‌های فاحش بین برآوردهای انجام شده در زمینه آمارهای فصلی حساب‌های ملی ایران که هر کدام تنها برای دوره خاصی برآورد شده اند، علی رغم نیاز بالای روش شبکه‌های عصبی مصنوعی به نمونه‌ای با داده‌های زیاد، ترجیح داده شده است که از آمارهای سالانه رسمی کشور در این مطالعه استفاده شود. برای جبران این معضل و افزایش نسبی طول سری‌های زمانی مورد استفاده، از مطالعه اخیر مؤسسه تحقیقات پولی و بانکی بانک مرکزی جمهوری اسلامی ایران در زمینه برآورد حساب‌های ملی ایران در دوره ۱۳۱۵-۱۳۳۷ که با توجه به اطلاعات

---

۲. البته، پژوهشهایی هم وجود دارند که در آنها شبکه‌های عصبی بهتر از مدلهای خطی عمل نکرده اند. از جمله این مقالات می توان به سوانسون و وایت(Swanson & White 1995,1997) ، چارچ و کارم (Church & Curram 1996) و استاک و واتسون1998) (Stock & Watson اشاره کرد.

Archive of SID

www.SID.ir

Archive of SID

موجود برای آن دوره و تقریباً به طور مستقیم محاسبه شده است، استفاده می‌کنیم[۱]. به این ترتیب، دوره مورد بررسی در این مطالعه ۱۳۱۵-۱۳۸۰ است که شامل ۶۶ داده سالانه است.

۵-۲. تخمین رگرسیون رشد

برای تخمین مدل خطی رشد اقتصادی در ایران از مطالعه قدیمی و مشیری[۲](۱۳۸۱) در خصوص تخمین رگرسیون رشد در ایران استفاده می کنیم. در مطالعه یادشده به طور مفصل به ادبیات و پژوهش‌های تجربی در زمینه رشد اقتصادی در دنیا و ایران پرداخته شده است. لذا در این مطالعه به این موارد اشاره نخواهد شد و نتایج مطالعه مذکور پذیرفته می شود.

در مطالعه مورد نظر، متغیرهای تأثیر گذار بر رشد اقتصادی در ایران در ۸ طبقه تقسیم بندی شده اند: ۱) تجارت، ۲) شرایط سیاسی، ۳) سرمایه گذاری و سرمایه فیزیکی، ۴) نیروی کار و جمعیت، ۵) سرمایه انسانی، ۶) متغیرهای پولی و اعتباری (گسترش مالی)، ۷) سطح قیمت‌ها و ۸) اندازه دولت. در نهایت، مدل رابطه معنی دار مثبتی بین نرخ سرمایه گذاری، رشد حجم تجارت، رشد سرمایه انسانی و رشد اقتصادی و رابطه منفی معنی داری بین تورم و بی ثباتی سیاسی و رشد اقتصادی در ایران را در دوره مورد بررسی گزارش می کند.

لازم به ذکر است که تخمین مدل رگرسیون رشد در مطالعه مورد نظر برای دوره ۱۳۱۵-۱۳۸۰ انجام شده است درصورتی که با توجه به هدف این مطالعه(مقایسه قدرت پیش‌بینی مدل خطی رگرسیون رشد و مدل غیر خطی شبکه عصبی مصنوعی) باید بین ۵ تا ۱۰ درصد از نمونه[۳] برای ارزیابی قدرت پیش‌بینی مدل در نظر گرفته شود(در این پژوهش دوره ۱۳۸۰-۱۳۷۴ به عنوان مجموعه آزمون انتخاب

---

۱. در این گزارش روش عمومی برآورد تولید ناخالص داخلی ایران در سالهای ۱۳۱۵ تا ۱۳۳۷ روش جزء پایه است که بر اساس آن نخست ارزش افزوده هر یک از اجزای تولید ناخالص داخلی برآورد گردیده و سپس از جمع زدن آنها تولید ناخالص داخلی به دست می آید. در برآورد هر یک از اجزای تولید نیز حتی الامکان روش مستقیم بکار گرفته شده است، مگر در برخی از فعالیت‌های بخش خدمات که از نتایج روابط رگرسیونی برای برآوردداستفاده شده است. برای مطالعه جزییات بیشتر به منبع اصلی به آدرس ذکر شده در فهرست منابع مراجعه کنید.

۲. در دست انتشار است.

۳. قاعده ای برای این منظور وجود ندارد. پژوهشگران پیشنهاد می کنند بین ۵ تا ۱۰ درصد از نمونه به مجموعه آزمون اختصاص داده شود.

Archive of SID

www.SID.ir

Archive of SID

شده است). بنابراین، متغیرهای مدل باید مجدداً در دوره جدید مورد آزمون ریشه واحد قرار گیرند و خود معادله نیز مجدداً برآورد شود.

۵-۳. بررسی ویژگی‌های سری‌های زمانی متغیرها

در مطالعه قدیمی و مشیری در زمینه برآورد رگرسیون رشد در ایران، از متغیر نرخ رشد تولید ناخالص داخلی سرانه به عنوان متغیر وابسته و از متغیرهای نرخ رشد حجم تجارت، نرخ رشد تعداد افراد با تحصیلات عالی در کل شاغلان (سرمایه انسانی)، نرخ تورم و نسبت تغییرات تشکیل سرمایه ثابت به تولید ناخالص داخلی(نرخ سرمایه گذاری) و یک متغیر مجازی که برای سال‌های جنگ جهانی دوم، ملی شدن صنعت نفت و جنگ تحمیلی عدد یک را اختیار می کند به عنوان متغیرهای مستقل استفاده شده و نتایج آزمون بهبود یافته دیکی-فولر ایستایی سری‌های مورد استفاده را نشان می دهد. البته، آزمون‌های مربوط در مطالعه مذکور برای دوره ۱۳۱۵-۱۳۸۰ انجام شده است. برای دوره ۱۳۱۵-۱۳۷۳ آزمون‌های لازم مجدداً انجام و ایستایی سری‌ها در طول این دوره نیز تأیید شد. بنابراین، امکان تخمین مدل از روش حداقل مربعات معمولی برای دوره مربوط نیز امکان پذیر است.

۵-۴. نتایج برآورد

با توجه به توضیحات مذکور، رگرسیون مطالعه قدیمی و مشیری(۱۳۸۱) برای دوره ۱۳۱۵-۱۳۷۳ مجدداً برآورد شد که نتایج آن به شرح زیر گزارش می شود:

$$RGDPL = \underset{(2.54)}{0.03} + \underset{(1.94)}{0.24}DIGDP + \underset{(1.80)}{0.09}RH + \underset{(4.23)}{0.16}RXM - \underset{(-1.61)}{0.06}P - \underset{(-2.02)}{0.04}D1$$

$$R^2 = 0.63 \qquad \overline{R}^2 = 0.60 \qquad F = 18.41$$

همان طور که مشاهده می شود کلیه ضرایب دارای علامت مورد انتظار هستند و متغیرهای نرخ رشد حجم تجارت و بی ثباتی در سطح ۵ درصد، متغیر تغییرات نرخ سرمایه گذاری در سطح ۶ درصد و

Archive of SID

www.SID.ir

Archive of SID

رشد سرمایه انسانی در سطح ۸ درصد معنی دارند. متغیر نرخ تورم نیز در سطح ۱۲ درصد معنی دار است(در منطقه مرزی قرار دارد)[۱].

در ادامه بحث ابتدا، یک شبکه عصبی مصنوعی با مجموعه متغیرهای ورودی مشابه با مدل رگرسیون رشد برآورد شده در بالا (رگرسورهای مشابه)، طراحی و تخمین زده می شود و سپس، توانایی دو مدل برای پیش‌بینی با استفاده از اطلاعات مجموعه آزمون(۱۳۷۴-۱۳۸۰) و معیارهای مرسوم مقایسه خواهند شد.

۶. طراحی شبکه عصبی مصنوعی برای رشد اقتصادی ایران

با توجه به اینکه مدل خطی برآورد شده بالا، با وجود ۵ رگرسور، تنها حدود ۶۳ درصد از تغییرات رشد تولید ناخالص داخلی سرانه در ایران را توضیح می دهد، احتمال وجود روابط غیر خطی در مدل بسیار زیاد است. از این رو در این قسمت، با استفاده از یک مدل شبکه عصبی مصنوعی پیش‌خور با یک لایه پنهان به مدل‌سازی پدیده رشد اقتصادی در ایران می پردازیم. همان‌طور که قبلاً تأکید شد، ثابت می شود یک مدل شبکه عصبی با یک لایه پنهان و تابع محرک لجستیک در لایه پنهان و یک تابع محرک خطی در لایه خروجی و تعداد کافی نرون‌ها در لایه پنهان قادر است هر رابطه غیر خطی را تقریب بزند. بنابراین، برای طراحی مدل رشد اقتصادی ایران یک شبکه عصبی با مجموعه متغیرهای ورودی گرفته شده از مدل رگرسیون بالا و یک لایه پنهان با تابع محرک تانژانت هیپربولیک[۲] در نظر

---

۱. مقدار آماره دوربین-واتسون معادله برآورد شده برابر۲/۵۹ است که بسیار نزدیک منطقه مرزی قرار دارد و نیاز است برای بررسی وجود خود همبستگی از آزمونهای دیگر استفاده کرد. اما، با توجه به اینکه هدف مطالعه حاضر مقایسه توان پیش‌بینی مدلهای رگرسیون با شبکه‌های عصبی است، فرض شد که معادله دارای خود همبستگی پیاپی است و مدل مجدداً، پس از رفع خود همبستگی احتمالی مورد برآورد قرار گرفت. اما، با توجه به اینکه قدرت پیش‌بینی معادله رگرسیون پس از رفع خود همبستگی احتمالی از معادله ابتدایی کمتر شد، لذا همان معادله اول برای مقایسه با مدل شبکه عصبی انتخاب شد.

۱. از آنجایی که متغیر خروجی ارزش‌های منفی را نیز اختیار می کند، از این تابع فعال سازی به جای تابع لجستیک استفاده شده است.

www.SID.ir

Archive of SID

می گیریم. در این مدل، متغیر خروجی رشد تولید ناخالص داخلی سرانه در ایران است و یک تابع فعال سازی تانژانت هیپربولیک[۱] برای واحد(نرون) خروجی پذیرفته می شود.

برای تعیین تعداد نرون‌های پنهان از فرمول‌های مرسوم[۲] استفاده می‌شود. بر اساس این فرمول‌ها تعداد نرون‌های پنهان باید برابر۲/۴۵ و ۱/۷ باشد[۳] . بنابراین، شبکه ای با دو نرون پنهان را درنظر می گیریم[۴] و با استفاده از الگوریتم لونبرگ – مارکوانت[۵] و ارزش‌های مختلف اولیه آنرا تخمین می زنیم[۶]. عمل تخمین زدن و یا به عبارت دقیق‌تر آموزش شبکه با ارزش‌های مختلف اولیه بارها برای فرار از گرفتار شدن در کمینه نسبی انجام می شود[۷].

یادآوری می شود برای دستیابی به قدرت تعمیم بالاتر شبکه ازهر دو روش منظم سازی بیزین و توقف زود هنگام استفاده شد که در نهایت، روش توقف زود هنگام نتیجه بهتری به دست داد. در این

---

۲. با توجه به اینکه ارزش‌های متغیر خروجی بین یک و منهای یک محدود است، استفاده از این تابع فعال سازی مناسب تر از سایر توابع است.

۳. هیچ پایه نظری برای انتخاب تعداد لایه‌ها و واحدهای پنهان وجود ندارد. در ادبیات شبکه‌های عصبی، برای پیدا کردن تعداد بهینه واحدهای پنهان چند فرمول پیشنهاد شده است که متأسفانه هیچ توافق عمومی نسبت به آنها وجود ندارد و استفاده از فرمول‌های مختلف نتایج مختلفی را به بار می آورد. برای نمونه می توان به دو فرمول $h = T/[5(r+v)]$ و $(\square\square)^{1/2}$ اشاره کرد که در آن h تعداد واحدهای پنهان، r تعداد متغیرهای ورودی، v تعداد متغیرهای خروجی وT تعداد مشاهدات در مجموعه آموزش است (Master,۱۹۹۴)

۱. فرمول اول، مجذور حاصل ضرب تعداد متغیرهای ورودی و خروجی و فرمول دوم، برابر تعداد مشاهدات مجموعه آموزش تقسیم بر مجموع تعداد متغیرهای ورودی و خروجی ضربدر ۵ است.

۲. البته، شبکه‌هایی با نرون‌های پنهان بیشتر نیز تخمین زده شد. اما، نتیجه بهتری به دست نیامد.

۳. البته، از الگوریتم‌های دیگر نیز می‌توان استفاده کرد. اما، سرعت و کارایی الگوریتم لونبرگ – مارکوانت برای تخمین شبکه مورد بررسی بالاتر است.

۴. برای این منظور، از جعبه ابزار شبکه‌های عصبی (Neural Network Toolbox) نرم افزار مطلب (MATLAB) استفاده شده است.

۵. یادآوری می شود نتیجه مربوط ممکن است با تکرار بیشتر و شروع آموزش شبکه با ارزشهای اولیه دیگری بهبود یابد. نتیجه به دست آمده بهترین نتیجه ممکن به وسیله نویسندگان است.

Archive of SID

www.SID.ir

Archive of SID

روش داده‌های دوره ۱۳۱۵-۱۳۶۶ به عنوان مجموعه آموزش، دوره ۱۳۶۷-۱۳۷۳ به عنوان مجموعه اعتبار و دوره ۱۳۷۴-۱۳۸۰ به عنوان مجموعه آزمون در نظر گرفته شد.

نتایج آزمون قدرت پیش‌بینی شبکه عصبی مصنوعی تخمین زده شده، با استفاده از ۵ معیار در مجموعه آزمون، با مدل خطی رگرسیون رشد برآورد شده در بخش قبل مقایسه شده است که نتایج آن در جدول (۱) خلاصه شده است.

جدول-۱. ارزیابی و مقایسه قدرت پیش‌بینی مدل شبکه‌های عصبی مصنوعی با مدل‌خطی‌رگرسیون

نوع مدل/نام معیار	RMSE	MAE	MAPE	TIC	CR
مدل ANN	۰/۰۲۳	۰/۰۱۹	۱۷۶/۶۷۶	۰/۴۹	۰/۵۷۱
مدل رگرسیون	۰/۰۳۵	۰/۰۳۱	۲۸۱/۷۵	۰/۶۰	۰/۷۱۴

همان طور که مشاهده می شود بر اساس کلیه معیارها(ریشه میانگین مربع خطا، میانگین قدرمطلق خطا، میانگین قدر مطلق درصد خطا، ضریب نابرابری تایل و نرخ اغتشاش) مدل شبکه‌های عصبی مصنوعی بطور قابل توجهی برتر از مدل رگرسیون است.

اما برای اینکه از نظر آماری نیز، فرضیه برابری صحت پیش‌بینی‌ها در دو مدل رقیب را نیز مورد آزمون قرار دهیم، از آماره دایبولد-ماریانو[1](۱۹۹۵) و آماره اصلاح شده آن به وسیله هاروی، لیبورن و نیوبلد[2](۱۹۹۷) استفاده می کنیم. در آزمون دایبولد-ماریانو ابتدا اختلاف خطاهای پیش‌بینی دو مدل رقیب به صورت زیر محاسبه می شود:

$$d_t = (y_t - \hat{y}_{tm_1})^2 - (y_t - \hat{y}_{tm_2})^2 \qquad t = 1,2,...,t_f$$

که در آن، y_t مقادیر واقعی متغیر مورد پیش‌بینی و $\hat{y}_{tm_1}$ و $\hat{y}_{tm_2}$ به ترتیب مقادیر پیش‌بینی شده مدل اول و دوم هستند. t نیز دوره مورد پیش‌بینی را نشان می دهد. آماره دایبولد-ماریانو به صورت زیر محاسبه می شود و برای نمونه‌های بزرگ دارای توزیع نرمال با میانگین صفر و واریانس یک است:

$$s = \frac{\bar{d}}{\sqrt{\phi}}$$

1. Diebold-Mariano Statistic
2. Harvey, Leybourne & Newbold

Archive of SID

www.SID.ir

Archive of SID

که در آن، $\bar{d}$ میانگین اختلاف خطاها در دوره مورد پیش‌بینی($\bar{d} = \frac{1}{t_f}\sum_{t=1}^{t_f} d_t$) و ϕ واریانس مجانبی آن است که می تواند به وسیله یک مجموع غیر وزنی از اتوکوواریانس d_t به صورت $\hat{\phi} = \frac{1}{t}\left(\hat{\gamma}_0 + 2\sum_{i=1}^{k-1}\hat{\gamma}_i\right)$ که در آن، k افق پیش‌بینی، $\hat{\gamma}_i$ تخمین iام اتوکوواریانس و $\hat{\gamma}_0$ تخمین واریانس است تخمین زده شود.

برای نمونه‌های کوچک هاروی، لیبورن و نیوبولد (۱۹۹۷) آماره اصلاح شده زیر را پیشنهاد می کنند:

$$s^* = s.\left(\frac{t_f + 1 - 2k + k(k-1)/t_f}{t_f}\right)^{1/2}$$

که در آن، آماره s^* با ارزشهای بحرانی توزیع t دانش آموز[1] با درجه آزادی $n-1$ (به این دلیل که واریانس d تخمین زده می شود)، مقایسه می شود.

از آنجایی که در این مطالعه، افق زمانی پیش‌بینی یک دوره بعد است، $\hat{\phi}$ به $\frac{1}{t}(\hat{\gamma}_0)$ کاهش می یابد.

اکنون، با استفاده از خطاهای پیش‌بینی دو مدل به محاسبه آماره هاروی، لیبورن و نیوبولد(به دلیل کوچک بودن نمونه مورد مطالعه) می پردازیم:

$$s^* = \frac{0.00071}{0.00028}\left(\frac{7 + 1 - 2(1) + 1(1-1)/7}{7}\right)^{1/2} = 2.35$$

همان طور که مشاهده می شود، مقدار آماره مذکور، از مقدار بحرانی توزیع t دانش آموز با درجه آزادی ۶ حتی در سطح معنی دار پنج درصد به میزان ۱/۹۴۳ بیشتر است. بنابراین، فرضیه صفر مبنی بر

1. Student t Distribution

Archive of SID

www.SID.ir

Archive of SID

برابر بودن صحت پیش‌بینی دو مدل رد می شود. به عبارت دقیق‌تر، مدل شبکه عصبی به طور معنی داری پیش‌بینی‌های دقیق‌تری از مدل رگرسیون خطی رشد در ایران ارایه می دهد.

۷. جمع بندی و نتیجه گیری

وجود روابط غیر خطی در پدیده‌های دنیای واقعی علاقه و انگیزه اقتصاددانان را به استفاده از مدل‌های غیر خطی در سال‌های اخیر افزایش داده است. در این میان، استفاده از متدولوژی شبکه‌های عصبی مصنوعی به عنوان یک ابزار قدرتمند تجزیه و تحلیل داده‌ها و مدل‌سازی روابط غیر خطی، در اقتصاد رو به گسترش است. ثابت می شود یک مدل شبکه عصبی ساده پیش‌خور با یک لایه پنهان و تابع فعال‌سازی لجستیک در نرون‌های لایه پنهان و یک تابع فعال سازی خطی در نرون خروجی قادر است هر رابطه غیر خطی را تقریب بزند. از این رو، شبکه‌های عصبی مصنوعی یک تقریب زننده جامع به حساب می آیند.

شبکه‌های عصبی مصنوعی کاربردهای مختلفی در اقتصاد دارند، اما، مهمترین کاربرد آنها در اقتصاد مدل‌سازی برای پیش‌بینی متغیرهای اقتصادی است.

در مطالعه حاضر ما ابتدا، با استفاده از یک مدل رگرسیون خطی، معادله رشد اقتصادی در ایران را با استفاده از داده‌های دوره ۱۳۱۵-۱۳۷۳ تخمین زدیم و با استفاده از آن، نرخ رشد اقتصادی در ایران را برای دوره ۱۳۸۰-۱۳۷۴ پیش‌بینی کردیم. سپس، یک مدل شبکه عصبی برای مدل‌سازی رشد اقتصادی در ایران، با متغیرهای ورودی مشابه مدل رگرسیون خطی(رگرسورهای مشابه) با یک لایه پنهان و تابع تانژانت هیپربولیک در لایه پنهان و نرون خروجی طراحی کردیم و با استفاده از الگوریتم پس انتشار خطا برای دوره مشابه مدل رگرسیون تخمین زدیم. در نهایت، قدرت پیش‌بینی دو مدل برای دوره ۱۳۸۰-۱۳۷۴ با استفاده از معیارهای مختلف با یکدیگر مقایسه شد. نتایج این ارزیابی کارایی بالاتر مدل شبکه عصبی مصنوعی را نسبت به مدل رگرسیون خطی نشان می دهد.

Archive of SID

www.SID.ir

Archive of SID

منابع

خاوری نژاد، ابوالفضل. (۱۳۸۰). *برآورد تولید ناخالص داخلی ایران*. (۱۳۳۷-۱۳۱۵)، مؤسسه تحقیقات پولی و بانکی بانک مرکزی جمهوری اسلامی ایران.

قدیمی، محمد رضا و مشیری، سعید. (۱۳۸۱). *بررسی عوامل رشد اقتصادی در ایران*، تخمین رگرسیون رشد. در دست انتشار.

Bosarge, W. E. (1993). Adaptive Processes to Exploit the Nonlinear Structure of Financial Market. In: R. R.Trippi and Turban(eds.), *Neural Networks in Finance and Investing*. Probus Publishing, p. 371-402.

Chatfield, Chris (1993). Neural Networks: Forecasting Breakthrough or Passing Fad?. *International Journal of Forecasting* 9, p. 1-3.

Church, Keith B. and Curram, Stephen P. (1996). Forecasting Consumers' Expenditure: A Comparison between Econometric and Neural Network Models. *International Journal of Forecasting* 12, p. 255-267.

Cybenko, G. (1989). Approximation by Superpositions of a Sigmoid Function. *Mathematics of Control, Signals, and Systems* 2, p. 303-314.

Diebold, F.X. and Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics* 13, p. 253-263.

Donaldson, Glen R. and Kamstra, Mark. (1996). Forecast Combining with Neural Networks. *Journal of Forecasting* 15, p. 49-61.

Fu, Jiarong (1998). A Neural Network Forecast of Economic Growth and Recession. *The Journal of Economics* XXIV, No. 1, p. 51-66.

Funahashi, K. (1989). On the Approximate Realization of Continuous Mappings by Neural Networks. *Neural Networks* 2, p. 183-192.

Haefke, C. and Helmenstein, C. (1996). Neural Networks in the Capital Markets: An Application to Index Forecasting. *computational Economics* 9, p. 37-50.

Hecht-Nielson, R. (1987). Kolmogorov's Mapping Neural Network Existence Theorem. 1st IEEE ICNN, Vol. 3, San Diago.

Harvey, D.I., Leybourne S.J. and Newbold, P. (1997). Tests for Forecast Encompassing. *Journal of Business and Economic Statistics* 16, p. 254-259.

Archive of SID

www.SID.ir

Archive of SID

Hill, Tim, Marquez, Leorey, O'Connor, Marcus and Remus, William. (1994). Artificial Neural Network Models for Forecasting and Decision Making. *International Journal of Forecasting* 10, p. 5-15.

Hiemstra, Y. (1996). Linear Regression versus Backpropagation Networks to Predict Quarterly Stock Market Excess Returns. *computational Economics* 9, p. 67-76.

Holden, K. , Peel, D. A. and Thompson, J. L. (1990). *Economic forecasting: An Introduction.* Cambridge: Cambridge University Press.

Hornik, K., Stinchcombe, M. and White, Halbert. (1989). Multi-Layer Feedforward Networks Are Universal Approximators. *Neural Networks* 2, p. 359-366.

Kohzadi, N., Boyd, M. S., Kaastra, I., Kermanshahi, B. S. and Scuse D. (1995). Neural Networks for Forecasting: An Introduction. *Canadian Journal of Agricultural Economics 43*, p. 463-474.

Kuan, Chung-Ming and White ،Halbert. (1994). Artificial Neural Networks: An Econometric Perspective. *Econometric Reviews* 13, p. 1-91.

Maasoumi, E., Khotanzad, A. and Abaye, A. (1994). Artificial Neural Networks for Some Macroeconomic Series: A First Report. *Econometric Reviews 13*, p. 105-122.

Masters, T. (1994). *Signal and Image Processing with Neural Network.* John Wiley & Sons Inc..

Moody, John, Levin, Uzi and Rehfuss, Steve. (1993). Predicting the U.S. Index of Industrial Production. *Neural Network World*, 3(6) , p. 791-794 in special issue: *Proceeding of Paralled Applications in Statistics and Economics 93*, Mirko Novak(ed.).

Moshiri, Saeed, Cameron, Norman and Scuse, David. (1999). Static, Dynamic, and Hybrid Neural Networks in Forecasting Inflation. *Computational Economics* 14, p. 219-235.

Moshiri, Saeed and Cameron, Norman. (2000). Neural Network Versus Econometric Models in Forecasting Inflation. *Journal of Forecasting* 19, p. 201-217.

Nelson, C. and Plosser, C. (1982). Trends and Random Walks io Macroeconomic Time Series: Some Evidence and Implications. *Journal of Monetary Economics 10*, p. 139-162.

Perron, p. (1989). The Great Crash, the Oil Price Shock and the Unit Root Hypothesis. *Econometrica 57*, p. 1361-1401.

Archive of SID

www.SID.ir

Archive of SID

Rech, G. (2002). Forecasting With Artificial Neural Network Models. *SSE/EFI Working paper Series in Economics and Finance*, No. 491.

Refenes, A. P. (1995). *Neural Networks in the Capital Markets*. Wiley.

Refenes, Apostolos Nicholas, Zapranis, Achieas and Francis, Gavin. (1995). Modelling Stock Returns in the Framework of APT: C Comparative Study with Regression Models. In: A. P. Refenes (ed.): *Neural Networks in the Capital Markets*. Wiley, p. 101-125.

Refenes, Apostolos Nicholas, Zapranis, Achieas and Francis, Gavin. (1994). Stock Performance Modelling Using Neural Networks: A Comparative Study with Regression Models. *Neural Networks*, Vol. 7, No. 2, p. 375-388.

Rumelhart, D. E., Hinton, G. E. and Williams, R. J. (1986). Learning Internal Representations by Error Propagation. *Parallel Distributed Processing 1*, p. 318-62.

Sargent, T. S. (1993). *Bounded Rationality in Macroeconomics*. Clarendon Press.

Sarle, Warren S . (1995). Stopped Training and Other Remedies for Overfitting. *Proceedings of the 27th Symposium on the Interface*.

Stock, James H. and Watson, Mark W. (1998). A comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series. *National Bureau of Economic Research Working Paper 6607*.

Swanson, Norman R. and White, Halbert. (1995). A Model-Selection Approach to Assessing the Information in the Term Structure Using Linear Models and Artificial Neural Networks. *Journal of Business and Economic Statistics*, Vol. 13, No. 3, p. 265-275.

Swanson, Norman R. and White, Halbert. (1997). A Model Selection Approach to Real-Time Macroeconomic Forecasting Using Linear Models and Artificial Neural Networks. *Review of Economics and Statistics* 79, p. 540-550.

Tkacz, Greg. (1999). *Neural Network Forecasts of Canadian GDP Growth Using Financial Variables.* Bank of Canada, mimeo.

Trippi, R. R. and Turban, E. (1990). Auto Learning Approaches for Building Expert System. *Computer and Operations Research* 17, p. 553-560.

Tsibouris, G. and Zeidenberg, M. (1995). Testing the Efficient Markets Hypotheses with Gradient Descent Algorithms. In: A. P. Refenes (ed.): *Neural Networks in the Capital Markets,* Wiley, p. 127-136.

Archive of SID

www.SID.ir

Archive of SID

White, H. (1988). Economic prediction using Neural Networks: The case of IBM Daily Stock Returns. *Proceeding of the IEEE International Conference on Neural Networks* II, p. 451-458.

White, H. (1992). *Estimation, Inference and Specification Analysis*. New York: Cambridge University Press.

Wong, F. S. (1990). Time Series Forecasting using Backpropagation Neural Networks. *Neurocomputing* 2, p. 147-159.

Verkooijen, William (1996). A Neural Network Approach to Long-Run Exchange Rate Prediction. *Computational Economics* 9, p. 51-65.

Archive of SID

www.SID.ir