

فصلنامه مدل‌سازی اقتصادسنجی (سال اول، شماره اول «پیاپی ۱»، تابستان ۱۳۹۳، صفحات ۳۳-۴۹)

پیش‌بینی و تحلیل نتایج اشتغال صنعتی در ایران با روش شبکه عصبی مصنوعی (ANN) و مدل ARIMA*

احمد جعفری صمیمی

استاد اقتصاد دانشکده علوم اقتصادی و اداری دانشگاه مازندران

jafarisa@umz.ac.ir

زهرا دهقان

دانشجوی کارشناسی ارشد اقتصاد دانشگاه مازندران (نویسنده مسئول)

zahra.dgn@gmail.com

تاریخ دریافت: ۱۳۹۲/۰۹/۲۵ تاریخ پذیرش: ۱۳۹۳/۰۱/۲۰

چکیده

صنعت یکی از بخش‌های مهم و اساسی اقتصاد و زمینه‌ساز رشد و توسعه صنعتی است. رشد و توسعه بخش صنعت، زمینه را برای رشد و توسعه سایر بخش‌ها از جمله کشاورزی، خدمات، حمل و نقل و انرژی فراهم می‌سازد. این بخش در فرایند توسعه نقش مهمی در ایجاد اشتغال دارد. با توجه به اهمیت پیش‌بینی در برنامه‌ریزی و سیاست‌های اقتصادی و اهمیت اشتغال در بخش صنعت، مطالعه حاضر به پیش‌بینی میزان اشتغال صنعتی ایران با استفاده از روش شبکه عصبی مصنوعی ANN و ARIMA پرداخته است. بدین منظور از داده‌های ۱۳۵۸-۱۳۹۰ استفاده شده است. برای بررسی دقت پیش‌بینی از ریشه میانگین مربعات خطا RMSE، میانگین مطلق درصد خطا RMSE و آماره‌ی U تایل استفاده شده است. نتایج تحقیق نشان می‌دهد که، شبکه عصبی پیشرو پس انتشار دارای قدرت بالایی در پیش‌بینی اشتغال صنعتی در ایران می‌باشد و نسبت به روش ARIMA دارای خطای کمتری است.

طبقه‌بندی JEL: C45, C22, E24, J21.

کلید واژه‌ها: اشتغال، پیش‌بینی، شبکه عصبی مصنوعی، ARIMA

این مقاله از پایان‌نامه کارشناسی ارشد زهرا دهقان تحت راهنمایی دکتر احمد جعفری صمیمی استخراج شده است.

۱. مقدمه

بیکاری و رشد اقتصادی پایین از مشکلات مهم و اساسی سیاست‌گذاران اقتصادی در هر کشور می‌باشد. با توجه به روند رو رشد جمعیت جهان و عرضه روز افزون نیروی کار، مقابله با بیکاری و ایجاد فرصت شغلی با وجود تلاش‌های بسیار همچنان از معضلات مهم اجتماعی، فرهنگی و اقتصادی در هر کشوری است. در بین بخش‌های اقتصادی، بخش صنعت و معدن، از بخش‌های مهم اقتصادی است و در فرایند رشد و توسعه، نقش مهمی در ایجاد اشتغال دارد.

پس از انقلاب صنعتی در اروپا و پیشرفت علوم، بخش صنعت یکی از بخش‌های اصلی و مهم اقتصاد کشورها شده است. این بخش علاوه بر جایگاه مهمی که در درآمد ملی دارد، نقش مهمی در اشتغال نیروی کار، درآمد و مصرف داشته و دارد. انتخاب استراتژی‌ها و سیاست‌های صنعتی در توسعه‌ی، بخش صنعت و به تبع آن در اشتغال‌زایی و تقاضای نیروی کار موثر بوده است (مفتح، ۱۳۸۶). اقتصاددانانی چون لوییس^۱ (۱۹۵۴)، کالدور^۲ (۱۹۶۶) بر آن بودند که گسترش بخش صنعت، نیروی محرک رشد اقتصادی و افزایش بهره‌وری در دیگر بخش‌ها است. تایلر^۳ (۱۹۷۴) نشان داده است که به هنگام تندتر شدن آهنگ رشد اقتصادی، بخش صنعت پیشرو است و زودتر از دیگر بخش‌ها توسعه می‌یابد. براساس مطالعه هریس^۴ (۲۰۰۱)، رشد سرمایه‌گذاری صنعتی در کشورهای روبه توسعه، به افزایش چشمگیر سهم نیروی کار شاغل در بخش صنعت انجامیده و آهنگ رشد و توسعه اقتصادی را تندتر کرده است. از این رو، بسیاری از کارشناسان رشد اشتغال صنعتی را کلید موفقیت کشورهای رو به توسعه بشمار آورده‌اند.

یکی از مسائل مهم در مدل‌سازی‌های اقتصادی پیش‌بینی مقادیر آینده یک متغیر است. دقت پیش‌بینی از مهمترین عوامل موثر در انتخاب روش پیش‌بینی است. با وجود روش‌های متعدد پیش‌بینی، اکثر محققان درصدد یافتن روش‌های هستند که به مقدار واقعی نزدیکتر بوده و خطای آن بسیار کم باشد. برخلاف بسیاری از متغیرهای اقتصادی، اشتغال و بخصوص اشتغال بخش صنعت از عرصه تحقیقات نوین پیش‌بینی به دور مانده، به گونه‌ای که مطالعه‌ای چندانی در زمینه پیش‌بینی اشتغال صنعتی در کشور

¹ Lewis

² Kaldor

³ Tyler

⁴ Harris

وجود ندارد. با توجه به اهمیت پیش‌بینی در برنامه‌ریزی‌ها و سیاست‌های اقتصادی و اهمیت اشتغال بخش صنعت، مطالعه حاضر به پیش‌بینی اشتغال صنعتی ایران به روش شبکه عصبی مصنوعی^۱ و $ARIMA$ پرداخته و دقت پیش‌بینی آنها مقایسه شده است.

۱. پیشینه تحقیق

در این قسمت به بیان دیدگاه‌ها و نظریه‌های مکاتب اقتصادی در مورد اشتغال و تقاضای نیروی کار پرداخته می‌شود. در تحلیل کینز درآمد ملی در شرایط رکود اقتصادی، در سطحی پایین‌تر از اشتغال کامل تعیین می‌شود. پایه‌های اساسی این تحلیل را سه اصل زیر تشکیل می‌دهند (کاکاوند قلعه‌نویی، ۱۳۸۷):

انعطاف‌پذیری قیمت‌ها با تشکیل انحصار از بین می‌رود و مسئله چسبندگی قیمت‌ها به خصوص قیمت نیروی کار (دستمزد) وجود دارد.

تعادل اشتغال کامل در بازار نیروی کار به وجود نمی‌آید و همواره بیکاری غیر ارادی جایگزین آن می‌شود.

تعادل بازار سرمایه به شکل برابری پس انداز و سرمایه‌گذاری بیانگر تعادل درآمد ملی است و به تعادل نرخ بهره منجر می‌شود که در این بازار نیز پدیده دام نقدینگی فرض انعطاف‌پذیری نرخ بهره را از بین می‌برد.

مکتب کلاسیک‌ها با تز لیبرالیسم اقتصادی و براساس فلسفه اصالت فرد در امور اقتصادی، در سال ۱۷۷۶ در انگلستان بنیان‌گذاری شد. آدام اسمیت با انتشار کتاب ثروت ملل بنیانگذار مکتب کلاسیک شناخته شد. طبق نظر اقتصاددانان مکتب کلاسیک، انعطاف‌پذیری قیمت‌ها، دستمزدها و میزان بهره، تعادل را در سه بازار نیروی کار (بازار اشتغال)، بازار سرمایه و بازار پول، در حالی که عرضه و تقاضا در اشتغال کامل هستند فراهم می‌سازد. به عبارتی در نظام کلاسیک به خاطر انعطاف‌پذیری دستمزد و قیمت از یک سو و انعطاف‌پذیری میزان بهره از سوی دیگر تعادل توأم با اشتغال کامل در بازار اشتغال و تولید برقرار شده و در نتیجه مجموع اقتصاد به طور دائم در تعادل همراه با اشتغال کامل قرار خواهد داشت. در دهه‌ی ۱۹۷۰ کلاسیک‌های جدید، اقتصاد کینز را به دلیل نداشتن پایه‌های قوی تئوریک زیر سوال بردند. آنها اعتقاد داشتند که نظریات اقتصاد کلان به دلیل فقدان زیر بنای خرد با مشکل مواجه بوده است و اقتصاد کلان

¹ Artificial Neural Network

² Autoregressive Integrated Moving Average

باید بر اساس پایه‌های تئوریک محکم اقتصاد خرد بنا شده باشد که اقتصاد کینز از این جنبه دچار ضعف بوده است. آنها معتقد بودند که انتظارات می‌تواند در تعیین رفتار پویا در اقتصاد کلان، نقش بسیار مهمی ایفا کند و مفهوم انتظارات عقلایی را مطرح نمودند. این گروه با قبول فرضیه تسویه دائم بازار و انتظارات عقلایی به وجود بیکاری ارادی در اقتصاد معتقد می‌باشند که این امر در حد نرخ طبیعی بیکاری می‌باشد (اشرفی پور و امیر عباسی، ۱۳۹۱: ۱۲۱). از اواسط دهه‌ی ۱۹۸۰ مکتب کینزین جدید به عنوان رقیب اصلی رویکرد کلاسیک جدید ظهور کرده است. مکتب کینزین‌های جدید با مکتب پولی ارتباط داشته و اساس نظریات این مکتب از ترکیب نظریات مکتب پولی و دو فرض رفتار عقلایی و رقابت ناقص تشکیل شده است (اشرفی پور و امیر عباسی، ۱۳۹۱: ۱۲۲). در مدل‌های کینزین جدید چسبندگی دستمزد و قیمت مورد تأکید قرار گرفته است. پول دیگر خنثی نبوده و کارساز بودن سیاستی آن دوباره احیا می‌شود. فیشر^۱ و فلیس^۲ و تایلور^۳ (۱۹۷۷) نشان دادند، نقش‌های اسمی تقاضا قابلیت ایجاد آثار حقیقی در مدل‌های دربردارنده انتظارات عقلایی را دارا می‌باشند. در این‌گونه الگوها، سیاست پولی نظام‌مند می‌تواند به تثبیت امور اقتصادی کمک کند. تعدیل تدریجی قیمت‌ها و دستمزدها بر این امر دلالت دارد که هر سیاست پولی ضد تورمی حتی اگر اطمینان بخش باشد، اگر به وسیله‌ی کارگزاران عقلایی پیش‌بینی شود، به یک رکود قابل توجه بر حسب تولید و اشتغال خواهد انجامید. افزون بر این، در برخی از تحلیل‌های کینزی جدید نرخ تعادلی بیکاری تحت تاثیر مسیر مربوط به نرخ بیکاری واقعی قرار می‌گیرد، به طوری که نرخ بیکاری تحت تاثیر تقاضای کل قرار می‌گیرد. تحلیل‌های کینزی جدید بر خلاف نگرش‌های کلاسیک جدید، برای وجود بیکاری غیر ارادی مبانی منطقی تدارک دیدند.

مطالعات زیادی در زمینه پیش‌بینی با استفاده از روش شبکه عصبی مصنوعی و ARIMA انجام گرفته است که ذیل به ذکر چند نمونه از این مطالعات که با پژوهش حاضر ارتباط دارند پرداخته شده است.

پوری و سویدمایر^۴ (۲۰۰۰)، برای ۵ کشور کالیفرنیا جنوبی به پیش‌بینی اشتغال صنعتی با استفاده از مدل‌های BVAR^۵، ARIM و VAR^۱ پرداختند و به این نتیجه رسیدند

^۱Fischer

^۲phelps

^۳tyler

^۴puri & Soydemir.

^۵Bayesian vector auto regressive

که مدل BVAR مدل بهتری از مدل نامحدود VAR و مدل تک متغیری ARIMA است. در افق‌های زمانی طولانی مدل BVAR به طور نسبی بهتر از مدل‌های دیگر است. کالوو^۲ (۲۰۰۵)، ۱۷ منطقه اسپانیا را بر اساس صنعتی‌بودن در ساختار اشتغال به سه گروه صنعتی شده، شبه صنعتی و کمتر صنعتی تقسیم کرده و رابطه بین رشد اشتغال صنعتی و سهم هر بنگاه و وجود نوآوری در صنایع را برای دوره زمانی ۲۰۰۲-۱۹۹۸ بررسی کردند. حیدر و ندیم حنیف^۳ (۲۰۰۹) با استفاده از داده‌های ۲۰۰۷-۱۹۹۳ و روش‌های اتورگرسیو مرتبه اول AR(1) و ARIMA و شبکه عصبی مصنوعی، به پیش‌بینی تورم در پاکستان پرداخته است. نتایج تحقیق نشان می‌دهد که شبکه عصبی مصنوعی دارای دقت بیشتری در پیش‌بینی دارد. سونگ وهان^۴ (۲۰۱۰)، در پژوهشی با استفاده از مدل‌های سری زمانی، اشتغال صنعتی در آلاسکا را پیش‌بینی نمودند و به این نتیجه رسیدند که در بلندمدت مدل VAR نسبت به مدل‌های دیگر بهتر عمل می‌کنند. پاتولی و همکاران^۵ (۲۰۱۱)، به پیش‌بینی اشتغال منطقه‌ای با مدل‌های توسعه یافته شبکه عصبی، با استفاده از داده‌های مربوط به دوره‌های ۲۰۰۴-۱۹۸۷ و ۲۰۰۴-۱۹۹۳ به ترتیب برای دو منطقه غرب و شرق آلمان پرداختند. آنها برای بررسی دقت پیش‌بینی از میانگین مربع خطا و میانگین خطای مطلق استفاده نمودند. نتایج نشان داد که ترکیب‌های مختلف از پارامترهای شبکه عصبی، کارایی و قدرت پیش‌بینی قابل توجهی را ارائه می‌دهد. مدل ARIMA برای سال‌های ۱۹۹۰ تا ۲۰۰۰ تا ARIMA(1,2,2) و برای ۲۰۰۱ تا ۲۰۰۹ تا ARIMA(2,2,1) و برای ۱۹۹۰ تا ۲۰۰۹ بشکل ARIMA(3,2,1) است. از مطالعات داخلی می‌توان به تحقیق بهزاد و ملک پور^۶ (۱۳۸۴)، اشاره نمود. آنها به منظور ارزیابی قابلیت اشتغال‌زایی در بخش صنعت ایران، به تعریف متغیرهایی که بر اشتغال صنعتی تاثیر گذار می‌باشند پرداختند. دلیری و خلیلیان^۷ (۱۳۸۴)، برای پیش‌بینی مقادیر رشد و تورم در بخش کشاورزی از مدل‌های هموارسازی نمایی حالت^۶ وینترز^۷ و مدل ARMA استفاده کردند و در ادامه به منظور مقایسه دقت پیش‌بینی مدل‌های خطی و غیرخطی، یک مدل شبکه عصبی مصنوعی

¹ vector auto regressive

² Calvo

³ Haidar and Nadeem-Hanif

⁴ Seung & Ahn

⁵ Patuelli & et al

⁶ Holt

⁷ Winters

مبتنی بر متغیرهای مدل رگرسیون و مدل ARIMA طراحی نمودند. مفتح (۱۳۸۶)، در مقاله خود به بررسی نقش صنعت در اشتغال‌زایی ایران پرداخته است. برای سنجش عملکرد بخش صنعت کشور از سه معیار، رشد صنعتی، رشد بهره‌وری سرانه نیروی کار و صادرات صنعتی استفاده کرد. نتایج تحقیق نشان می‌دهد نسبت اشتغال در گروه صنعت و معدن از ۳۱/۲ درصد در سال ۱۳۵۶ به ۲۲/۸ درصد در سال ۱۳۷۸ کاهش یافته است. پدram و حبیبی فرد (۱۳۸۷)، در تحقیق خود به منظور شناسایی عوامل موثر بر تقاضای نیروی کار در بخش صنعت و بررسی رابطه بلند مدت میان تقاضای نیروی کار و عوامل موثر بر آن به برآورد تقاضای نیروی کار در بخش صنعت ایران طی سال‌های ۸۳-۱۳۵۰ با استفاده از تکنیک هم‌انباشتگی جوهانسون پرداختند. کهنسال (۱۳۹۱) با استفاده از سه روش ANN, ARIMA و هموارسازی حالت-وینترز به پیش‌بینی قیمت تخم مرغ ایران با استفاده از داده‌های ۱۳۸۶-۱۳۸۰ پرداخته است. برای بررسی مانایی از آزمون دیکی فولر تعمیم یافته استفاده شده است. برای انتخاب بهترین مدل از معیار میانگین مربع خطا MSE و معیار میانگین قدرمطلق انحراف MAD و آماره u تایل استفاده شده است. نتایج برآورد نشان می‌دهد روش شبکه عصبی مصنوعی ANN نسبت به روش هموارسازی و ARIMA خطای کمتری دارد.

۳. روش تحقیق

در این تحقیق از دو روش شبکه عصبی مصنوعی و روش ARIMA برای پیش‌بینی اشتغال صنعتی ایران استفاده شده است که در این قسمت به توضیح این دو روش پرداخته شده است.

۳-۱. مدل ARIMA

اگر یک سری زمانی پس از d مرتبه تفاضل‌گیری مرتبه‌ی اول ساکن شود و سپس توسط فرآیند $ARMA(p, q)$ مدل‌سازی شود، سری زمانی اصلی سری زمانی خود رگرسیونی میانگین متحرک انباشه $ARIMA(p, d, q)$ است که در آن p تعداد جملات خودرگرسیو، d تعداد دفعات تفاضل‌گیری مرتبه اول برای ساکن شدن سری زمانی و q تعداد جملات میانگین متحرک می‌باشد. مدل ARIMA، مقادیر آینده متغیر به عنوان تابعی خطی، از مشاهدات گذشته و خطاهای تصادفی فرض می‌شوند، یعنی فرآیند اساسی که سری زمانی را تولید می‌نماید بدین صورت می‌باشد:

$$y_t = \theta_0 + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (1)$$

آن y_t مقادیر واقعی، ε_t خطاهای تصادفی در دوره t ، و پارامترهای p و q مرتبه مدل می‌باشند. فرض می‌شود خطای تصادفی ε_t مستقل و دارای توزیع یکنواخت با میانگین صفر و واریانس ثابت σ^2 است. در حالتی که $q=0$ باشد رابطه (۱)، به یک معادله اتو رگرسیو از درجه p می‌باشد و وقتی $p=0$ باشد مدل به یک مدل میانگین متحرک از درجه q تبدیل می‌شود (خاشعی و بیجاری، ۱۳۸۶: ۸۸).

متدولوژی باکس_جنکینز دارای چهار مرحله است:

-تشخیص (شناسایی)،

-تخمین.

-کنترل تشخیص (بازبینی).

-پیش‌بینی.

مرحله اول، تشخیصی (شناسایی): در این مرحله مقادیر مربوط به وقفه‌های جملات خودرگرسیون، تعداد دفعات تفاضل‌گیری برای مانا شدن و مرتبه مدل تعیین می‌شود. به عبارتی در این مرحله مقادیر واقعی p, d, q تعیین می‌شود. بدین منظور از آزمون دیکی - فولر تعمیم یافته و ابزار نمودار همبستگی و همبستگی جزئی استفاده می‌شود. مرحله دوم، تخمین: این مرحله عمدتاً وابسته به تشخیص مدل در مرحله اول است. در این مرحله پارامترهای مدل به روش حداقل مربعات معمولی برآورد می‌شوند.

مرحله سوم، کنترل تشخیصی (بازبینی): پس از انتخاب یک مدل خاص و تخمین پارامترهای آن به این سوال پاسخ داده می‌شود که آیا مدل انتخابی، داده‌ها را به خوبی برازش می‌کند یا خیر؟ باکس جنکینز دو روش را مطرح نموده است: یکی برآورد مدل با افزودن مرتبه AR و MA می‌باشد. اگر مدل شناسایی شده در مرحله اول کفایت کند در این صورت افزودن مرتبه AR و MA هیچ کمکی نمی‌کند و معنی‌دار نخواهد بود. روش دیگر بازبینی باقیمانده‌ها است. بازبینی باقیمانده‌ها به معنی کنترل باقیمانده‌ها است که آیا شواهدی دال بر وابستگی خطی وجود دارد یا نه (سوری، ۱۳۸۹: ۱۷۳). که می‌بایست باقی مانده‌های حاصل از این مدل نوفه سفید باشد. در صورتی که باقی مانده‌ها نوفه سفید باشند می‌توان مدل انتخابی را به عنوان یک برازش مناسب پذیرفت، در غیر این صورت مدل رد شده و مراحل فرآیند دوباره تکرار می‌شود.

مرحله چهارم، پیش‌بینی: در این مرحله، از مدل نهایی برای پیش‌بینی سری زمانی استفاده می‌شود. در بسیاری از موارد پیش‌بینی‌های حاصل از روش $ARIMA$ که

به ویژه برای کوتاه مدت استفاده می شود، بیش از روش مدل سازی سنتی اقتصادسنجی قابل اعتماد و اتکا است.

۳-۲. شبکه عصبی مصنوعی

استفاده از شبکه عصبی برای پیش بینی و مدل سازی در اقتصاد کلان و اقتصادسنجی در دهه ی ۹۰ آغاز شد و مطالعات متعددی نیز در زمینه استفاده از شبکه های عصبی مصنوعی در پیش بینی متغیرهای مختلف اقتصادی بصورت گرفت. مهم ترین ویژگی مدل های ANN آزادی آنها از فروض آماری مربوط به متغیرها، استفاده از روش های محاسباتی موازی و غیر خطی بودن آنهاست. مدل های شبکه عصبی بسیار متنوع هستند ولی ساختار آنها شبیه به هم می باشد. شبکه عصبی معمولاً از سه لایه تشکیل شده است:

لایه ورودی^۱ داده ها را در شبکه توزیع می کند.

لایه پنهان^۲: داده ها را پردازش می کند.

لایه خروجی^۳: نتایج را به ازای ورودی های مشخص استخراج می کند.

لایه ورودی فقط اطلاعات را دریافت می کند و مشابه متغیر مستقل عمل می کند. لذا تعداد نرون های لایه ورودی به تعداد متغیرهای مستقل بستگی دارد. لایه خروجی نیز همانند متغیر وابسته عمل کرده و تعداد نرون های آن بستگی به تعداد متغیرهای وابسته دارد. محققین از روابط مختلفی از جمله $n/2$, $2n$, $2n+1$ که n تعداد نرون های لایه ورودی می باشند برای تعیین تعداد نرون های لایه مخفی استفاده نموده اند. نقش نرون ها در شبکه عصبی، پردازش اطلاعات است و این امر در شبکه عصبی به وسیله یک پردازش ریاضی که تابع فعال سازی است، انجام می شود. توابع فعال سازی، می توانند خطی و یا غیر خطی باشند (طرازکار و نجفی، ۱۰: ۱۳۸۵). مهمترین توابع فعال سازی غیر خطی مورد استفاده تابع سیگموئید^۴ و تابع تانژانت هیپربولیک^۵ می باشند. تابع سیگموئید سیگموئید که مقدار آن در محدوده صفر و یک است. به شکل زیر می باشد (سام دلیری و خلیلیان، ۱۳۸۴: ۲۰۰):

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2)$$

¹ Input Layer

² Hidden layer

³ Output layer

⁴ Sigmoid Function

⁵ Hyperbolic Tangent

اگر متغیر پیش‌بینی شده، بتواند مقادیر منفی را نیز اخذ کند، بهتر است از تابع فعال‌سازی تانژانت هیپربولیک استفاده شود. مقدار این تابع در محدوده‌ی ۱ و -۱ تغییر می‌کند (سام دلیری و خلیلیان، ۱۳۸۴: ۲۰۰):

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

شبکه‌های عصبی مصنوعی دارای مدل‌های مختلفی هستند که مبتنی بر جهت ورود اطلاعات و پردازش آنها به انواع مختلف تقسیم می‌شوند. برای انجام این تحقیق از شبکه‌های عصبی پیشرو پس انتشار^۱، استفاده گردید. در شبکه پیشرو، گره‌ها در لایه‌های متوالی قرار گرفته‌اند و ارتباط آنها یک طرفه است و زمانی که یک الگوی ورودی به شبکه اعمال می‌شود، اولین لایه مقادیر خروجی‌اش را محاسبه کرده و در اختیار لایه بعدی قرار می‌دهد. لایه بعدی این مقادیر را به عنوان ورودی دریافت کرده و مقادیر خروجی‌اش را به لایه بعدی منتقل می‌کند و هر فقط به گره‌های بعدی سیگنال منتقل می‌کند. واژه پس انتشار نیز به معنای این است که خطاها به سمت عقب در شبکه تغذیه می‌شوند تا وزن‌ها را اصلاح کنند و پس از آن مجدداً ورودی، مسیر پیشروی خود تا خروجی را تکرار کند. روش پس انتشار خطا از روش‌های با ناظر است. به این مفهوم که نمونه‌های ورودی برچسب خورده‌اند و خروجی مورد انتظار هر یک از آنها، از پیش معلوم است. لذا خروجی شبکه با این خروجی‌های ایده‌آل مقایسه شده و خطای شبکه محاسبه می‌گردد. در این الگوریتم ابتدا فرض بر این است که وزن‌های شبکه به طور تصادفی انتخاب شده‌اند. هر گام خروجی شبکه محاسبه شده و بر حسب میزان اختلاف آن با خروجی مطلوب، وزن‌ها تصحیح می‌گردند تا در نهایت این خطا مینیمم شود (نوائی، ۱۰۴: ۱۳۹۱).

در نهایت به منظور مقایسه قدرت پیش‌بینی و انتخاب بهترین روش پیش‌بینی، از معیارهای مختلفی استفاده می‌شود. در این تحقیق از معیارهای ریشه میانگین خطا^۲ RMSE، ریشه میانگین مطلق درصد خطا^۳ MPE و آماره‌ی^۴ U تایل که صورت زیر تعریف و اندازه‌گیری می‌شود استفاده شده است:

^۱ Feed Forward backpropagation Neural Network

^۲ Root Mean Square Error

^۳ Mean Absolute Percentage Error

^۴ The U Sttisti

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (F_t - A_t)^2}{n}} \quad (4)$$

$$MPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (5)$$

$$U = \frac{\sqrt{\frac{\sum_{t=1}^n (F_t - A_t)^2}{n}}}{\sqrt{\frac{\sum_{t=1}^n (A_{t-1} - A_t)^2}{n}}} \quad (6)$$

۴. یافته‌های پژوهش

۴-۱. نتایج مدل ARIMA

در بحث مدل‌های ARIMA مانایی و نامانایی، یک سری می‌تواند تاثیر جدی بر رفتار و خواص آن دارد. استفاده از داده‌های نامانا می‌تواند موجب رگرسیون کاذب شود. بنابراین باید قبل از مدل‌سازی مانایی سری بررسی شود. در این پژوهش برای شناسایی مانایی و نامانایی از آزمون دیکی- فولر و یا دیکی- فولر تعمیم یافته استفاده می‌شود. نتایج آزمون ریشه واحد برای اشتغال صنعتی در جدول (۱) آورده شده است.

جدول ۱: نتایج آزمون دیکی فولر

متغیر	مقدار بحرانی	آماره آزمون	prob
EMP	-۲/۹۶۰۴	-۰/۳۶۲	۰/۹۰۳
DEMP	-۲/۹۶۰۴	-۶/۱۵۹	۰/۰۰۰

نتایج آزمون ریشه واحد در جدول (۱) نشان می‌دهد که فرضیه H_0 در سطح ۵ درصد دارای ریشه واحد بوده و مانا نمی‌باشد. بنابراین جهت مانا شدن تفاضل مرتبه اول گرفته می‌شود. نتیجه ریشه واحد تفاضل مرتبه اول، حاکی از عدم وجود ریشه واحد در سطح ۵ درصد است بنابراین متغیر مورد نظر مانا می‌باشد. پس از بررسی مانایی، باید تعداد وقفه‌های خودرگرسیونی و میانگین متحرک را با استفاده از معیارهای تعیین وقفه آکائیک و شوارتز _ بیزین تعیین شود. جدول (۲) مقادیر آکائیک و جدول (۳) مقادیر شوارتز _ بیزین مدل‌های برآورد شده با وقفه‌های مختلفی از خودرگرسیونی و میانگین متحرک را نشان می‌دهد. تعداد جملات میانگین متحرک q و تعداد جملات خودرگرسیونی p است.

جدول ۲: مقادیر آکائیک در الگوی برآورد شده $ARIMA(p,q)$

q \ p	۰	۱	۲
۰	۱۲/۸۳۸	۱۳/۳۷۳	۱۳/۴۵۸
۱	۱۳/۷۶۶	۱۳/۴۸۸	۱۳/۲۳۸
۲	۱۳/۴۴۹	۱۲/۸۹۰	۱۲/۹۱۷

جدول ۳: مقادیر شوارتز - بیزین در الگوی برآورد شده $ARIMA(p,q)$

q \ p	۰	۱	۲
۰	۱۲/۷۹۱	۱۳/۲۸۰	۱۳/۳۱۸
۱	۱۳/۲۸۴	۱۳/۳۳۷	۱۳/۲۳۸
۲	۱۳/۳۱۲	۱۲/۷۰۵	۱۲/۸۳۳

همانطور که ملاحظه می‌شود در این مورد هر دو معیار آکائیک و شوارتز-بیزین الگوی $ARMA(2,2)$ را معرفی می‌نمایند. زیرا کمترین مقدار را نسبت به سایر وقفه‌ها دارند و چون مرتبه تفاضل‌گیری $d=1$ می‌باشد، مدل قابل تخمین به فرم $ARIMA(2,1,2)$ می‌باشد.

پس از بدست آوردن مرتبه مدل $ARIMA$ در مرحله دوم متدولوژی باکس - جنکینز مدل انتخاب شده تخمین زده می‌شود. برآورد مدل عمدتاً وابسته به تشخیص مدل در مرحله اول است. در این پژوهش مدل با روش OLS برآورد می‌شود. نتایج تخمین ضرایب مدل $ARIMA(2,1,2)$ در جدول (۴) آمده است.

جدول ۴: تخمین مدل $ARIMA(2,1,2)$

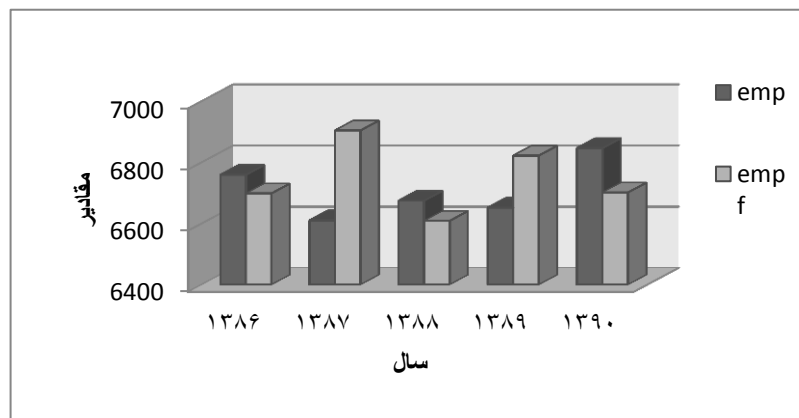
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	۱۷۷/۶۷۶	۲۲/۰۴۸	۸/۰۵۸	۰/۰۰۰۰
AR(۲)	۰/۶۰۸	۰/۱۴۰۰	۴/۳۴۵	۰/۰۰۰۲
MA(2)	-۰/۹۰۹۶۱	۰/۰۴۳۵	-۲۰/۹۰۳۷	۰/۰۰۰۰
R-squared	۰/۲۶۷	Akaike info criterion		۱۳/۲۷۷
Durbin-Watson	۱/۹۲۶۳	Schwarz criterion		۱۳/۴۱۷
F-statistic	۳/۵۱۸۳	Prob(F-statistic)		۰/۰۴۳۸

میزان Prob که سطح معنی‌دار بودن ضرایب را نشان می‌دهد کمتر از $۰/۰۵$ است در نتیجه آماره t استیوودنت، حاکی از معنی‌دار بودن، ضرایب خودرگرسیون $AR(۲)$ و وقفه میانگین متحرک $MA(۲)$ در سطح معنی داری ۵ درصد است. همچنین سطح

معنی دار بودن F یعنی $\text{Prob}(F\text{-statistic})$ برابر $0/0438$ و چون کمتر از $0/05$ است بنابراین آماره F آن نیز معنی دار است. همگی این نتایج نشان از خوبی برازش است. به منظور بررسی درستی مدل انتخاب شده، بعد از تخمین به کنترل آن پرداخته می شود برای کنترل از نمودارهای همبستگی پسماند^۱ استفاده می شود. با توجه به نمودارهای همبستگی پسماندها نمودارهای خودهمبستگی^۲ (ACF) و خودهمبستگی جزئی^۳ (PACF) حاکی از آن است که با تعیین طول وقفه مناسب هر یک از وقفه ها در فاصله اطمینان یک انحراف معیار قرار گرفته اند. آماره Q لجانگ - باکس برای همه وقفه ها بی معنی است. بنابراین، می توان نتیجه گرفت الگوی برآورد شده الگوی مناسبی است و می توان با استفاده از این مدل به پیش بینی برای دوره های فراتر از دوره تخمین اقدام نمود.

در مرحله آخر مدل ARIMA پیش بینی اشتغال صنعتی با استفاده از اطلاعات موجود گرفته است در این مدل از سال ۱۳۵۸ تا ۱۳۸۵ برای برآورد مدل و از سال ۱۳۸۶-۱۳۹۰ برای مقایسه بین مقادیر واقعی و پیش بینی شده توسط مدل ARIMA استفاده شده است. این نتایج در نمودار آن در شکل (۱) آورده شده است. emp مقادیر واقعی و empf مقادیر پیش بینی شده را نشان می دهد.

شکل ۱: نمودار مقایسه بین مقادیر واقعی و پیش بینی شده توسط مدل ARIMA



¹ Residual Test

² Autocorrelation

³ Partial Autocorrelation

نتایج مقادیر پیش‌بینی شده اشتغال صنعتی ایران به روش ARIMA در جدول (۵) آورده شده است.

جدول ۵: نتایج پیش‌بینی به روش ARIMA

سال	۱۳۹۱	۱۳۹۲
ARIMA	۷۰۲۶۶۶۹۵	۷۲۲۵۷۲۷

۴-۱. نتایج شبکه عصبی مصنوعی

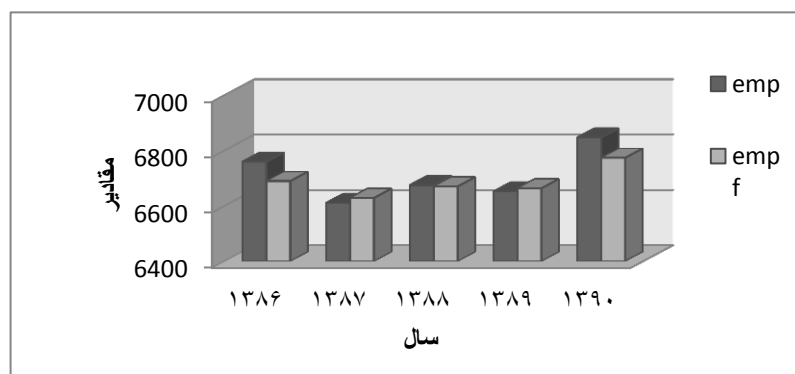
در این پژوهش از شبکه‌های عصبی پیشرو پس انتشار، استفاده می‌شود. در این نوع شبکه خروجی هر لایه، ورودی لایه بعدی است. پس از انتخاب نوع شبکه و آزمون باید ورودی‌های مدل یا همان واحدهای لایه ورودی، یا به عبارت دیگر متغیرهای توضیحی مدل انتخاب شود. شبکه‌ای که برای این تحقیق انتخاب شد دارای سه لایه است: لایه ورودی، لایه مخفی و لایه خروجی. تعداد بیشتر لایه‌ها شبکه را پیچیده می‌کند و تنها در صورتی که خروجی‌های بهتری از شبکه دریافت شود باید اقدام به افزایش لایه‌ها نمود. در تعیین تعداد نرون‌های لایه مخفی، باید به این نکته توجه شود که تعداد زیاد نرون‌ها در لایه مخفی، گرچه باعث پایین آمدن خطای آموزش می‌شود، اما از طرف دیگر، خطای آزمون را افزایش می‌دهد. بنابراین معمولاً از تعداد کم نرون شروع می‌کنند و در صورت بهبود جواب‌ها نرون‌ها افزایش می‌یابد. در این تحقیق، تعداد نرون‌ها از ۱ تا ۴۰ نرون در لایه مخفی آزمایش شده است. خطای آموزش در نرون ۶ به کمترین حد رسیده است. داده‌ها به دو گروه آموزش و آزمون تقسیم شده‌اند. از داده‌های سال‌های ۱۳۵۸ تا ۱۳۸۵ به عنوان گروه آموزش و از داده‌های سال‌های ۱۳۸۶-۱۳۹۰ به عنوان گروه آزمون استفاده شده است. به منظور انتخاب تابع فعال‌سازی سه تابع فعال‌سازی خطی، تانژانت هیپربولیک و سیگموئید و برای انتخاب الگوریتم آموزش نیز سه الگوریتم نیوتن، شبه نیوتن و لوتبرگ _ مارکوانت مورد آزمایش قرار گرفته است. نتایج حاصل از آزمایش اجزا مختلف در جدول (۶) آمده است.

نتایج آزمایشات مختلف نشان می‌دهد که بهترین نتایج با شبکه‌ای بدست می‌آید که دارای تابع فعال‌سازی لایه خروجی سیگموئید و تابع فعال‌سازی در لایه میانی تانژانت هیپربولیک است و با الگوریتم آموزش لوتبرگ مارکوانت می‌باشد. نتایج مقایسه مقادیر واقعی و پیش‌بینی گروه آزمون در شکل (۲) آمده است.

جدول ۶: نتایج حاصل از آزمایش اجزای مختلف شبکه عصبی

تایع فعال سازی لایه میانی	تایع فعال سازی لایه خروجی	الگوریتم آموزش	میانگین خطای پیش بینی	
تانزانت هیپربولیک	خطی	نیوتن	۰/۰۳۶۲	
		شبه نیوتن	۰/۰۴۴۲	
		لونبرگ مارکوانت	۰/۱۱۸	
	تانزانت هیپربولیک	تانزانت هیپربولیک	نیوتن	۰/۰۱۰۴
			شبه نیوتن	۰/۰۱۳۳
			لونبرگ مارکوانت	۰/۰۱۶۱
	سیگموئید	سیگموئید	نیوتن	۰/۰۱۷۴
			شبه نیوتن	۰/۰۱۷۵
			لونبرگ مارکوانت	۰/۰۱۱۰
سیگموئید	خطی	نیوتن	۰/۰۴۶۵	
		شبه نیوتن	۰/۰۵۰۲	
		لونبرگ مارکوانت	۰/۰۸۷۱	
	تانزانت هیپربولیک هیپربولیک	تانزانت هیپربولیک هیپربولیک	نیوتن	۰/۰۲۱۳
			شبه نیوتن	۰/۰۱۸۸
			لونبرگ مارکوانت	۰/۰۳۱۳
	سیگموئید	سیگموئید	نیوتن	۰/۰۲۶۸
			شبه نیوتن	۰/۰۲۸۳
			لونبرگ مارکوانت	۰/۰۷۰۲

شکل ۲: نتایج مقایسه مقادیر واقعی و پیش بینی به روش شبکه عصبی



emp مقادیر واقعی و empf مقادیر پیش‌بینی شده را نشان می‌دهد. مقادیر پیش‌بینی شده به روش شبکه عصبی مصنوعی در جدول (۷) آورده شده است.

جدول ۷: نتایج پیش‌بینی به روش شبکه

سال	۱۳۹۱	۱۳۹۲
ANN	۶۷۴۴۱۳۲	۶۶۹۴۰۰۵

میزان خطای پیش‌بینی شده، در دو روش، با ریشه میانگین مربعات خطا (RMSE)، درصد میانگین قدرمطلق خطا (MAPE) و آماره U تایل سنجیده شده است نتایج محاسبات در جدول (۸) آمده است.

جدول ۸: مقایسه خطای روش شبکه و ARIMA

RMSE	MAPE	U
۶۸/۰۹	۰/۰۱۸ ANN	٪۵۶
۴۶/۳۳	۰/۰۰۴	٪۳۴

همان‌گونه که مشاهده می‌شود، خطای پیش‌بینی در روش شبکه عصبی بر اساس هر یک از معیارها کمتر بوده، بنابراین شبکه عصبی مصنوعی نسبت به روش ARIMA در پیش‌بینی اشتغال صنعتی ایران عملکرد بهتری دارد.

۵. نتیجه‌گیری

این پژوهش به مقایسه روش شبکه عصبی مصنوعی ANN و ARIMA برای پیش‌بینی میزان اشتغال صنعتی ایران با استفاده از داده‌های ۱۳۵۸-۱۳۹۰ پرداخته است. شبکه بکارگرفته شده شبکه عصبی پیشرو پس انتشار است. به منظور انتخاب بهترین ساختار شبکه، شبکه‌هایی با توابع فعال‌سازی مختلف و الگوریتم‌های آموزش مختلف مورد آزمایش قرار گرفت و در نهایت، شبکه‌ای، با تابع فعال‌سازی خروجی سیگموئید و تابع فعال‌سازی میانی تانژانت هیپربولیک و الگوریتم آموزش لونیبرگ مارکوانت به عنوان بهترین شبکه برای پیش‌بینی انتخاب شد. نتایج تحقیق نشان می‌دهد شبکه عصبی پیشرو پس انتشار، دارای قدرت بالایی در پیش‌بینی اشتغال صنعتی ایران دارد و نسبت به روش ARIMA دارای عملکرد بهتری است. با توجه به نتایج این تحقیق که حاکی از برتری شبکه عصبی نسبت به روش ARIMA است و نیز با در نظر گرفتن توانایی بروز رسانی سریع شبکه‌های عصبی پس از ایجاد تغییرات سری‌زمانی مورد پیش‌بینی، پیشنهاد می‌شود در مدل‌سازی و پیش‌بینی متغیرهایی که رفتار

غیرخطی، از خود نشان می دهند، از مدل های غیرخطی نظیر شبکه های عصبی استفاده شود، زیرا این امر سبب کاهش خطا و بهبود نتایج پیش بینی خواهد شد.

فهرست منابع:

- اشرفی پور، محمدعلی و امیرعباسی، زهرا (۱۳۹۱)، بررسی ارتباط حجم نقدینگی و سطح اشتغال بخش صنعت در ایران. فصلنامه تحقیقات توسعه اقتصادی، ۷: ۱۱۱-۱۳۲.
- بانک مرکزی جمهوری اسلامی ایران اداره حساب های اقتصادی، ۱۳۸۱.
- بهزاد، رامین و ملک پور، کیانا (۱۳۸۴)، قابلیت اشتغال زایی بخش صنعت، نشریه کار و جامعه، ۶۱: ۱۶-۲۱.
- پدرام، مهدی و حبیبی فر، مریم (۱۳۸۸)، بررسی رابطه بلندمدت تقاضای نیروی کار و عوامل موثر بر آن در بخش صنعت ایران از طریق آزمون همجمعی جوهانسون، پژوهشنامه اقتصادی، ۱۴۱: ۱-۲۱.
- خاشعی، مهدی و بیجاری، مهدی (۱۳۸۷)، بهبود عملکرد پیش بینی های مالی با ترکیب مدل های خطی و غیر خطی خودرگرسیون میانگین متحرک انباشته و شبکه عصبی مصنوعی. فصلنامه پژوهش های اقتصادی، ۲: ۸۳-۱۰۰.
- زراءنژاد، منصور، فقه مجیدی، علی و رضایی، روح الله (۱۳۸۷)، پیش بینی نرخ ارز با استفاده از شبکه های عصبی مصنوعی و مدل ARIMA، فصلنامه اقتصاد مقداری، ۴: ۱۰۷-۱۳۰.
- سام دلیری، احمد و خلیلان، صادق (۱۳۸۴)، پیش بینی نرخ رشد و نرخ تورم در بخش کشاورزی ایران. مجله تحقیقات اقتصادی، ۷۴: ۱۸۳-۲۱۵.
- سوری، علی (۱۳۸۹)، اقتصادسنجی، نشر فرهنگ شناسی، چاپ اول.
- طرازکار، محمدحسین و نجفی، بهالدین (۱۳۸۵)، پیش بینی میزان صادرات پسته ی ایران: کاربرد شبکه عصبی مصنوعی، پژوهشنامه بازرگانی، ۳۹: ۷۲-۹۶.
- کاکاوند قلعه نویی، فرج (۱۳۸۷)، تئوری و سیاست اقتصاد کلان چاپ اول، ۴۶۷-۴۶۸.
- کهنسال، محمدرضا و پرمه، زورار (۱۳۹۱)، پیش بینی قیمت تخم مرغ با استفاده از ARIMA، شبکه عصبی مصنوعی و هموارسازی هالت وینترز. فصلنامه پژوهشنامه بازرگانی، ۶۲: ۴۹-۷۲.
- مرکز آمار ایران، نتایج آمارگیری نیروی کار، ۱۳۹۰.
- نوائی، نبوشا (۱۳۹۱)، تشخیص صحت امضای دستنویس بصورت برون خط، پایان نامه کارشناسی ارشد، دانشکده برق و رباتیک، دانشگاه صنعتی شاهرود.

Calvo, J. L. (2005), Industrial Employment Growth in Spanish Region the

Role Played By Size, Innovation and Spatial Aspects, Retrieved from: <http://www.Sre.wu.wien.ac.at/ersa/ersaconfs/ersa05/51.pdf>.

Harris, D. (2001), Industrial Employment, Investment and Equipment and Economic Growth, Center for Economic Policy Research, Economic Development and cultural change, 49: 867-881

Hidar, A and M. nadeem-Hanif (2009), Inflation Forecasting In Pakistan Using Artificial Neural Network, Pakistan Economic and Social Review, 47(1):123-138.

Kaldor, N. (1966), Causes Of The Slow Rates Of Economic Growth in the UK, Cambridge University Press.

Lewis, W. A. (1954), Economic Development with Unlimited Supplies of Labor, the Manchester School, 91-139.

Patuelli, R., Reggiani, A., Nijkamp, P. and N. Schanne (2011), Journal of Geographical Systems, 13: 67-85.

Puri, A and G. Soydemir (2000), Forecasting Industrial Employment Figures in Southern California: A Bayesian Vector Autoregressive Model. the Annals of Regional Science, 34: 503 – 514.

Seung, C, K and S. K. Ahn (2010), Forecasting Industry Employment for a Resource Based Economy Using Bayesian Vector Autoregressive Models. the Review of Regional Studies, 2: 181 – 196.

Tyler, W. (1974), Labor Absorption With Import-Substituting Industrialization: An Examination of Elasticities of Substitution in the Brazilian Manufacturing Sector, Oxford Economic Papers, New Series, 1: 93-103.

Forecasting Industrial Employment in Iran using Artificial Neural Network Method and ARIMA Model

Ahmad jafari samimi

Professor Of Economics, Faculty Of Economics Mazandaran
University, Babolsar, Iran

Jafarisa@umz.ac.ir

Zahra dehghan

M.A Of Economics, Mazandaran University, Babolsar Iran

Zahra.dgn@gmail.com

Abstract

Industry is one of the important and fundamental parts of economic and a ground for economic growth and development. Development and growth in industry section provide the ground for growth and development of other section such as agricultural, service, transport and energy. This section plays an important role in providing job in development process. According to the importance of forecasting in economic planning and policy-making and the importance of employment in industrial section, the present study dealt with forecasted number of industrial employment of Iran ANN and ARIMA artificial neural network method. For this reason 1358-1390 data were used. examine the validity of the research Mean Absolute Percentage Error MAPE, Root Mean Square Error RMSE and Theil U Ststistic were used. The research results show that back-propagation neural network has a high power in forecasting industrial employment in Iran and has lesser error in comparison to ARIMA method.

JEL Classification: C45, C22, E24, J21

Keywords: Employment, Forecasting, Artificial Neural Network, ARIMA