

پیش‌بینی میزان رشد ازدواج و طلاق در استان خراسان

جنوبی به کمک شبکه عصبی*

زهرة نیکوروش^۱

طاهره نیکوروش^۲

علیرضا مهرآور^۳

چکیده

پیش‌بینی‌های جمعیت در زمینه‌های مختلف از جمله ارزیابی نیازهای ملی یا منطقه‌ای در زمینه‌ی مشاغل جدید، معلم‌ها، مدارس، پزشکان، پرستاران، سکونت‌گاه‌های شهری و یا شناخت نیازهای غذایی شمار افرادی که باید به آن‌ها خدماتی ارائه شود، مفید واقع می‌شوند. بنابراین، پیش‌بینی‌های جمعیت نقطه‌ی شروع پیش‌بینی‌هایی است که مورد نیاز آینده هستند. امروزه تجزیه و تحلیل سری‌های زمانی به صورت گسترده در علوم مختلف مورد استفاده واقع می‌شود. از طرفی، پیش‌بینی رفتار در سیستم‌های پیچیده با استفاده از روش‌های کلاسیک و تحلیلی امری دشوار و در برخی موارد غیرممکن به نظر می‌رسد، از اینرو روش‌های غیرکلاسیک که از ویژگی‌هایی همچون هوشمندی برخوردار هستند، استفاده می‌شود. شبکه‌های عصبی، یکی از روش‌هایی است که در مسائلی مانند الگوسازی، شناخت الگو، خوشه‌بندی و پیش‌بینی مورد استفاده قرار می‌گیرد. در امر پیش‌بینی جمعیت وقایع چهارگانه حیاتی از جمله، تولد، وفات، ازدواج و طلاق دخیل هستند، که در این مقاله سعی شده با استفاده از تجزیه و تحلیل داده‌های سری زمانی با استفاده از شبکه‌های عصبی میزان ازدواج و طلاق را در سال‌های اخیر بررسی کرده و به مدل‌سازی و پیش‌بینی این میزان در استان خراسان جنوبی در سال‌های آتی بپردازیم.

واژگان کلیدی: نرخ رشد ازدواج، نرخ رشد طلاق، پیش‌بینی، شبکه عصبی، آمار حیاتی، استان خراسان جنوبی.

* برگرفته از طرح پژوهشی با عنوان "بررسی و تحلیل آمارهای وقایع حیاتی در استان خراسان جنوبی"

^۱ عضو هیات علمی دانشگاه صنعتی بیرجند و نویسنده مسئول: nikooravesh@birjandut.ac.ir

^۲ کارشناس ارشد علوم کامپیوتر- سیستم‌های هوشمند

^۳ کارشناس ارشد صنایع-مدیریت سیستم و بهره‌وری

مقدمه

گرچه هنوز بیش از ۵۰ سال از تولد روش‌های محاسباتی مبنی بر شبکه‌های عصبی مصنوعی نمی‌گذرد، لکن این شبکه‌ها به دلیل ویژگی‌هایی همچون پردازش موازی، هوشمندی و انعطاف‌پذیری جایگاه چشم‌گیری در مسائل پیچیده از قبیل شناخت الگو، خوشه‌بندی، مدل‌سازی، تخمین و شناسایی و پیش‌بینی برای خود باز کرده است. شبکه‌های عصبی از جنبه‌های توپولوژی، ساختاری و روش‌های یادگیری به انواع مختلفی تقسیم می‌شوند و هر یک در کاربردهای خاصی عملکرد مناسبی از خود نشان می‌دهند. شبکه عصبی چندلایه پرسپترون با روش انتشار وارون یکی از مهمترین شبکه‌های کاربردی است، که در مباحث نظری اثبات شده شبکه MLP در صورت انتخاب صحیح ساختار مناسب داخلی، قادر است هر گونه سیستم غیرخطی را مدل کرده و شبیه سازی کند. در این مقاله به کمک شبکه‌های عصبی که در رده روش‌های هوشمند برای حل مسائل قرار می‌گیرند، به پیش‌بینی آینده در سریهای زمانی پرداخته می‌شود. برخی از پژوهش‌های انجام شده ۱۳۳۵ تا در بحث جمعیت، سری‌های زمانی و شبکه‌های عصبی به شرح زیراند.

میرزائی و شمس قهفرخی (۱۳۸۶) جمعیت‌شناسی سالمندان در ایران را بر اساس سرشماری‌های سال‌های ۱۳۸۵ مورد مطالعه قرار دادند. شکوری (۱۳۸۸) طی یک طرح پژوهشی وضعیت جمعیت ایران را از منظر تعهدات صندوق بازنشستگی کشوری مورد بررسی اجمالی قرار داد و تسریع در تعدیل پارامترهای سیستم بازنشستگی و نهایتاً استقرار سیستم بازنشستگی چند لایه را پیشنهاد نمود. زمانی (۱۳۸۸) با بهره‌گیری از منابع آماری از قبیل آمارهای توصیفی، نتایج سرشماری نفوس و مسکن، با تمرکز بر مهاجرت‌های استانی به مطالعه جمعیت در برنامه‌ریزی آمایش سرزمین جهت دستیابی به توسعه متوازن در سطح ملی پرداخت. محمودی و دیگران (۱۳۹۱) با رویکردی میان رشته‌ای، متغیر راهبردی رشد جمعیت بر اساس متغیرهای اجتماعی و اقتصادی موثر بر آن، در قالب سه سناریوی اقتصادی، یعنی ادامه روند کنونی، تاثیر اجرای قانون هدفمند کردن یارانه ها و تحقیق اهداف چشم‌انداز ۱۴۰۴، مبتنی بر نظریه اینگلهارت، الگوسازی و تا افق ۱۴۰۴ پیش‌بینی کردند. آن‌ها نشان دادند که در هر یک از سناریوهای اقتصادی، رشد جمعیت، روند کاهنده را ادامه خواهد داد که نظریه نوسازی و تحول فرهنگی را درباره خیزش ارزش‌های فرامادی تأیید می‌کند.

اصغری اسکویی (۱۳۸۱) جهت پیش‌بینی سری‌های زمانی داده‌های اقتصادی از شبکه‌های عصبی استفاده کرده است و رابطه عوامل مختلف ساختاری، روش‌های مختلف یادگیری شبکه‌های عصبی را در فرایند پیش‌بینی، مورد ارزیابی و بررسی قرار داد. حقیقت منفرد و دیگران (۱۳۹۱) به مقایسه مدل‌های شبکه عصبی و سری زمانی در پیش‌بینی قیمت شاخص سهام پرداختند، آن‌ها سه مدل

شبکه عصبی و یک مدل سری زمانی را مورد مطالعه قرار دادند. شاری و دیگران (۲۰۱۴) به مدل سازی و پیش‌بینی تعداد زائران وارد شده به کشور عربستان سعودی بر اساس شبکه‌های عصبی توسط روش بیزین پرداختند. دیاکنسکو (۲۰۰۸) نیز به کمک شبکه‌های عصبی و با استفاده از روش NARX سری‌های زمانی منظم را مورد بررسی قرار داد.

مفاهیم و تعاریف مورد نیاز

در این بخش به بیان تعاریف و مفاهیم مورد نیاز می‌پردازیم. ابتدا تعاریفی را در مورد آمار حیاتی و سپس به مباحث درونیایی و سیستم‌های هوشمند و کاربردهای آنها در مساله پیش‌بینی خواهیم پرداخت.

آمار حیاتی، جمع‌آوری آمار درباره وقایع حیاتی در طول زندگی یک فرد و ویژگی‌های مرتبط به وقایع، فرد و افراد ذریع را شامل می‌گردد. آمار حیاتی، اطاعات مهم و اساسی را درباره جمعیت یک کشور در اختیار قرار می‌دهد. برای اهداف آماری، وقایع حیاتی وقایعی هستند که به مرگ و زندگی افراد و خانواده‌ها و جایگاه ثبتي آنها مربوط می‌شوند. ثبتي وقایع حیاتی به زندگی و مرگ مربوط می‌شود و شامل: تولدهای زنده، مرگ و مرگ‌های جنینی می‌گردد. وقایع دوگانه، وقایعی هستند که به طور همزمان در زندگی دو فرد رخ می‌دهند و نمی‌توانند بار دیگر در زندگی یکی از این دو فرد بدون تغییر قبلی در وضعیت او اتفاق بیفتند. این وقایع: ازدواج، همزیستی ثبتي شده، متارکه، طلاق، انحلال قانونی، همزیستی ثبتي شده و فسخ ازدواج را شامل می‌گردند. نهایتاً، وقایع خانوادگی عمودی، وقایعی هستند که اولاد را در برمی‌گیرند و شامل: فرزند خواندگی، مشروعیت و به رسمیت شناختن آن می‌گردند. فهرستی از هر واقعه که اطلاعات مربوط به آن باید برای اهداف آمار حیاتی جمع‌آوری گردند و تعریف پیشنهادی برای آن، مستقیماً در پایین آورده شده‌اند:

اهمیت پیش‌بینی آمار حیاتی

با توجه با کاربردهایی که آمارهای حیاتی دارند بررسی و پیش‌بینی وضعیت آینده آنها بسیار حائز اهمیت است. کسب دانش درباره اندازه و ویژگی‌های جمعیت یک کشور در بازه‌های زمانی مشخص، پیش‌نیاز برنامه‌ریزی اجتماعی، اقتصادی و تصمیم‌گیری‌های آگاهانه است. آمار حیاتی و تحلیل‌ها و تفسیرهای بعدی آنها برای تعیین هدف‌ها و ارزیابی برنامه‌های اجتماعی و اقتصادی از جمله پایش برنامه‌های مداخلات سلامت، جمعیت و سنجش شاخص‌های مهم جمعیت شناختی در سطوح زندگی یا کیفیت زندگی همچون: امید زندگی در هنگام تولد و نرخ مرگ و میر نوزادان همگی ضروری هستند. آمار حیاتی، همچنین برای برنامه‌ریزی، پایش و ارزیابی برنامه‌های

گوناگون، ارزشمند هستند؛ برنامه‌هایی همچون: برنامه مراقبت‌های بهداشتی اولیه، تغذیه، آموزش، مسکن عمومی و غیره. در میان کاربردهای جمعیت‌شناختی آمار حیاتی می‌توان از تهیه تخمین‌ها و پیش‌بینی‌های جمعیتی، مطالعه مرگ و میر، باروری، نرخ ازدواج و ساخت جداول زندگی نام برد.

منابع جمع‌آوری آمار

منبع اصلی آمار حیاتی، اسناد وقایع حیاتی است که از ثبت احوال استخراج می‌شوند و به جمع‌آوری مستمر اطلاعات درباره همه وقایع حیاتی مربوطه اطلاق می‌گردد که در چارچوب مرزهای یک کشور یا ناحیه کاملاً مشخص در کشوری رخ می‌دهند.

منابع آمار تکمیلی مانند: سرشماری‌های جمعیتی و سرشماری عمیق خانوارها هم، برای ارزیابی داده‌های ثبت احوال و غنا بخشیدن به آن‌ها و جمع‌آوری اطلاعات در مورد رویه‌های جمعیت‌شناختی و اپیدمیولوژیکی استفاده شده‌اند که اطلاعات بدست آمده از ثبت احوال را تکمیل می‌کنند.

منابع دیگر درون نظام آمار حیاتی شامل پرسش‌هایی خاص درباره باروری و مرگ و میر که به سرشماری‌های جمعیتی اضافه شده‌اند، سرشماری‌های نمونه خانوار، اسناد حیاتی ثبت نمونه و اسناد بهداشتی می‌گردند. از طریق استفاده از این منابع اطلاعاتی به همراه استفاده از تکنیک‌های غیرمستقیم تخمین‌های جمعیت‌شناسی، برخی کشورها شاخص‌های آماری خاصی را در اختیار گرفته‌اند که برای مقاصد برنامه‌ریزی مخصوصاً در سطح ملی لازمند. با این حال، هیچ جایگزینی برای وجود اطلاعات مستمر در مورد وقایع حیاتی بر اساس ثبت احوال آن‌ها وجود ندارد.

آمار حیاتی ترجیحاً از نظام ثبت احوال گرفته می‌شوند، زیرا این منبع ایده‌آلی برای استخراج اطلاعات دقیق، جامع، به‌موقع و مستمر درباره وقایع حیاتی است. به علاوه، آمار حیاتی استخراجی از نظام ثبت احوال و (دفاتر ثبت جمعیت) می‌توانند جریان آماری سالیانه از کوچکترین تقسیمات ثبت احوال را در برگیرند که هیچ نظام گردآوری داده‌ای، نمی‌تواند آن‌ها را فراهم سازد. وقتی ورودی ثبت احوال وجود ندارد یا ناقص است، بعضی کشورها ممکن است مجبور شوند به یک سرشماری جمعیتی یا مطالعه نمونه خانوار متوسل شوند تا آمار حیاتی لازم را از طریق سؤالات گذشته‌نگر درباره باروری، مرگ و میر و ازدواج تخمین بزنند. آمارهای باروری، مرگ و میر و ازدواج را ممکن است با تاسیس نواحی نمونه ثبت، جمع‌آوری کرد. در برخی کشورها، آمار حیاتی مورد نیاز برای برنامه‌ریزی به این منابع، دیگر اطلاعات به همراه کاربرد فنون غیرمستقیم تخمین‌های جمعیت‌شناختی وابسته هستند. باید تأکید شود که هرچند سرشماری‌های جمعیتی، بررسی نمونه‌ها و نمونه ثبت‌ها ممکن است تخمینی از سطوح باروری، مرگ و میر، مرگ و میر جنینی،

ازدواج و طلاق و در مورد نمونه ثبت، تخمینی از وقایع مرگ و میر به دلیل فوت به دست دهند، ولیکن آن‌ها جایگزینی برای نظام ثبت احوال نیستند، چرا که نمی‌توانند چنین جزئیاتی را درباره بازه‌های منظم سالانه و با پوششی همگانی ارائه دهند. یکی از مزایای آمار حیاتی تولید شده از ثبت احوال، پوشش جغرافیایی و گروه‌های جمعیتی کوچک است. داده‌های کافی از ثبت احوال که از سطح بالایی از پوشش در سطح ملی برخوردارند هم این ظرفیت را دارند که برآوردهای مجزا در سطوح منطقه‌ای را ممکن سازند. بنابراین اطلاعات ارزشمندی را برای برنامه‌ریزی منطقه‌ای و اختصاص درست منابع در حوزه‌هایی همچون آموزش، مراقبت‌های بهداشتی و امنیت اجتماعی در سطح مناسب اداری فراهم آورند.

مادامی که ثبت احوال، همگانی، مستمر و دائمی باشد، آمار حیاتی استخراج شده از ثبت احوال تنها منبع ملی نشانگر اطلاعات درباره مرگ و میر حاصل از فوت است. چنین اطلاعاتی برای ارزیابی و پایش وضعیت سلامت یک جمعیت و برنامه‌ریزی برای مداخلات مناسب بهداشتی ارزشمند هستند. ثبت به موقع فوت‌ها بر اساس علت، می‌تواند دیدی زود هنگام نسبت به روند شیوع بیماری‌ها به دست دهد و در نتیجه به طراحی راهبردهای پیشگیری و مداخله کمک کند. داده‌های قابل استناد و به موقع درباره علت مرگ هم این امکان را فراهم می‌کنند تا هشدارهای سریعی در حوزه بهداشت عمومی درباره مرگ در اثر بیماری‌های نادر داده شوند. اطلاعات مربوط به الگوهای نامعمول مرگ و علت‌های مرگ ممکن است به مقامات بهداشتی نشان دهند که نیاز به مداخله است.

پیش بینی روند رشد

برنامه‌ریزی کلید موفقیت هر برنامه آماری است. داده‌های حیاتی و فرمی که داده‌ها را ثبت می‌کند نوع آماری را تعیین می‌کند که می‌تواند پردازش شود. فارغ از روش پردازش، آمار گردآوری و جدول‌بندی شده نمی‌توانند از داده‌هایی که این آمار از آن‌ها استنتاج شده‌اند دقیق‌تر و کامل‌تر باشند.

برنامه پردازش آماری باید چندین هدف را در نظر بگیرد. تضمین این که اطلاعات مورد نیاز کاربران اصلی جمع‌آوری شوند. از آنجایی که پاسخ به نیازهای همه کاربران ممکن نیست، ضرورت دارد که اولویت‌های کاربران مشخص شوند و تلاش برای تحقق همه آن‌ها مهمترین امر دانسته شود. و اینکه برنامه‌ریزی بلند مدت لازم است زیرا تدارکات برای انجام یک برنامه آماری در یک سال مشخص معمولاً از چندین سال قبل زمان لازم دارد.

فرآیندی را که از طریق آن جمعیت‌شناسان به پیش‌بینی درباره جمعیت آینده می‌پردازند، پیش‌بینی جمعیت می‌گویند. یکی از روش‌های انجام پیش‌بینی روش ریاضی (تصادد حسابی) است. از دیگر روش‌ها می‌توان به کاربرد درونیایی و برونیایی داده‌ها اشاره کرد که بر اساس توصیف داده‌ها با یک فرمول ریاضی می‌توان به پیش‌بینی ادامه روند رشد پرداخت. در سال‌های اخیر استفاده از روش‌های هوشمند در بسیاری از زمینه‌های مختلف علوم رایج شده است. ما در این بخش ابتدا به اختصار به معرفی روش ریاضی و روش درونیایی و برونیایی داده‌ها پرداخته و ویژگی‌های مثبت و منفی آنها را بیان می‌کنیم.

روش ریاضی

در روش ریاضی به کمک تصاعد حسابی، تصور می‌شود تغییرات جمعیت طبق تصاعد حسابی است یعنی تغییر در تعداد جمعیت طبق سال‌های سرشماری ثابت می‌باشد. از فرمول زیر می‌توان به پیش‌بینی جمعیت در سال مورد نظر دست یافت.

$$p_{t+n} = p_t(1+r)^n \quad (1)$$

که در آن p_{t+n} : جمعیت در سرشماری دوم، p_t : جمعیت در سرشماری اول، n فاصله بین دو سرشماری و r نرخ رشد سالانه جمعیت است.

حجم جمعیت در سال مورد نظر برای پیش‌بینی در واقع متأثر از عواملی چون باروری، مرگ و میر، مهاجرت و تغییرات مربوط به ساختار سنی و جنسی جمعیت است. از آن جا که اثر مجموعه این عوامل را به ندرت می‌توان به طور یکجا با هم ارزیابی کرد از اهمیت و دقت این روش کاسته می‌شود. بنابراین در استفاده از میزان رشدی که به این صورت به دست آمده باشد، باید کاملاً احتیاط نمود و هرگز نباید از آن برای پیش‌بینی در دوره‌های دراز مدت استفاده کرد.

در رویکرد ریاضی پیش‌بینی جمعیت به طور آشکار هیچگونه تلاشی برای مدل‌سازی عناصر تغییرات جمعیت صورت نمی‌گیرد. به طور عادی این روش در مورد کل جمعیت به کار می‌رود، اما امکان کاربرد آن برای پیش‌بینی زیر گروه‌های جمعیتی وجود ندارد.

روش ریاضی سریع ساده و نیاز به داده‌های کمی دارد. چنان چه جمعیت مورد بررسی ثابت یا شبه ثابت باشد، برای پیش‌بینی جمعیت به تفکیک سن نیز می‌توان از این روش استفاده کرد. از لحاظ نظری مشخصه جمعیت ثابت این است که در طول زمان ساختار جمعیت بدون تغییر باقی بماند. به هر حال، در طول زمان نوساناتی در میزانهای روند باروری، مرگ و میر و مهاجرت رخ می‌دهد که خود منجر به تغییر ساختار جمعیت می‌گردد.

روش درونیابی و برونیابی داده‌ها

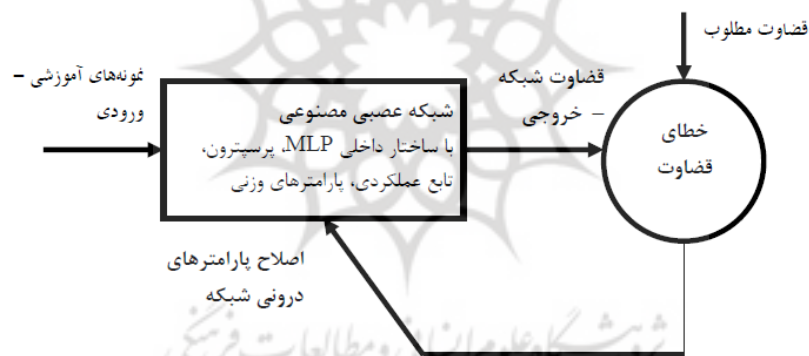
درونیابی جبری بر نقاط اندازه‌گیری شده متکی است و برای درونیابی از توابع ریاضی استفاده می‌کند. این درونیابی فرض می‌کند که تخمین پدیده مورد نظر قطعی انجام می‌شود و با خطا مواجه نیست، بنابراین درونیابی جبری یک درونیابی غیر احتمالی است. این درونیابی به کمک اندازه‌گیری‌های نقطه‌ای دقیق انجام می‌شود و اگر فرض شود که این اندازه‌گیری‌ها بدون خطا انجام گرفته‌اند آنگاه تابع درونیابی به گونه‌ای تعیین می‌شود که مقادیر برآوردی دقیقاً با مقادیر اندازه‌گیری شده برابر شوند. این گونه درونیابی دقیق است و مقدار برآوردی تنها تابعی از ساختار مکانی متغیر مورد نظر است و در آن نشانی از تغییرات تصادفی وجود ندارد.

درونیابی جبری بر دو گروه همگانی و محلی تقسیم می‌شود. در درونیابی همگانی برای برآورد نقاط مجهول از داده‌های همه نقاط اندازه‌گیری شده استفاده می‌شود، اما در مدل محلی از داده‌های همه نقاط برداشت شده برای برآورد مقدار نقطه مجهول استفاده نمی‌شود. در این روش ممکن است، تعداد نقاط محدودی که تا نقطه مورد نظر کمترین فاصله را دارند در درونیابی بکار گرفته شوند یا از داده‌هایی استفاده شود که تا نقطه مورد نظر کمترین فاصله را دارند. به بیانی دیگر در این روش نقاط همسایه بر اساس تعداد یا فاصله تعریف می‌شوند. اگر تغییرات مکانی متغیر زیاد باشد مدل‌های محلی نتایج بهتری بدست می‌دهد. با این حال تعداد نقاط همسایه که برای برآورد در یک نقطه معین بکار گرفته می‌شوند در کیفیت نتایج تأثیر بسزایی دارد. از آنجا که اندازه‌گیری‌ها همواره با مقداری خطا همراهند و برآوردها نمی‌توانند دقیقتر از اندازه‌گیری‌ها باشند، لذا روش‌های درونیابی جبری تنها زمانی سودمند خواهند بود که مقدار خطای اندازه‌گیری به اندازه کافی کوچک باشد.

شبکه‌های عصبی

شبکه عصبی مصنوعی، که مدل‌های ساده شده‌ی سیستم نرون‌های بیولوژیک هستند، سیستم پردازش توزیع موازی هستند و از تعداد فراوانی المان‌های عصبی محاسباتی متصل به هم تشکیل شده است که قابلیت یادگیری و کسب دانش را دارند و در حل مسائلی مانند شناسایی و تشخیص الگو، مدل‌سازی سری‌های زمانی و پیش‌بینی بر اساس آن‌ها، مدل‌سازی رگرسیون و به ویژه مدل‌سازی الگوهای غیرخطی، و بسیاری زمینه‌های دیگر کاربرد دارند. شبکه‌های عصبی مصنوعی قادرند روابط پیچیده غیرخطی بین متغیرهای ورودی و خروجی را در یک مسئله یاد بگیرند، بدون اینکه دانشی قبلی نسبت به آن مسئله داشته باشند.

شبکه عصبی در کاربرد همانند یک ذهن زنده عمل می‌کند. به این معنا که از مشاهدات انتزاعی خود به قضاوت می‌پردازد. لذا، شبکه عصبی مدتی را صرف آموزش کرده و سپس به صورت عملیاتی به کار گرفته می‌شود. شبکه عصبی آنچه را مشاهده می‌کند در قالب پارامترهای درونی خود به خاطر می‌سپارد. آنچه در ذهن شبکه عصبی نگهداری می‌شود، نه تک تک مشاهدات بلکه روش و برداشت کلی از مشاهدات است. به همین دلیل است که بعضا شبکه عصبی در رویارویی مجدد با نمونه‌های آموزشی چه بسا با خطای قابل اغماضی عکس‌العمل نشان می‌دهد. اما این استواری و ثبات در عمل را دارد که در برخورد با عموم نمونه‌های مشابه، عملکردی مناسب و همراه با خطای قابل چشم‌پوشی داشته باشد. یک شبکه عصبی چه بسا در حل یک مسئله بسیار خوب و در حل مسئله‌ای دیگر بد عمل کند. لذا، انتخاب ساختار شبکه، متناسب با موضوع مسئله از اهمیت بسیار بالایی برخوردار است. چرا که میزان یادگیری شبکه عصبی نیز از شرایط اولیه آموزش شبکه تاثیر می‌پذیرد. انتخاب مناسب مقادیر اولیه پارامترهای شبکه در نتیجه آموزش آن، بسیار موثر خواهد بود. شکل ۱ نمایی از آموزش شبکه عصبی را نشان می‌دهد.

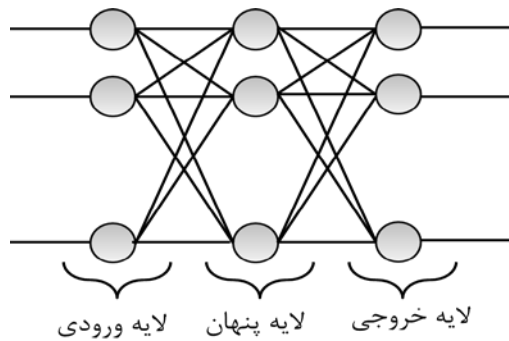


شکل(۱): نمایی از نحوی آموزش یک شبکه عصبی

ساختار شبکه عصبی شامل تعدادی پرسپترون با تابع عملکردی مشخص است که در لایه‌های مجزا قرار دارند. هر پرسپترون، به واسطه ضرایب وزنی خود خروجی تمامی پرسپترون‌های لایه قبلی را تجمیع کرده و از طریق تابع عملکردی به لایه بعدی ارسال می‌کند. یکی از ساده‌ترین و در عین حال کارآمدترین مدل شبکه عصبی، پرسپترون چندلایه^۱ (MLP) می‌باشد که از یک لایه ورودی، یک یا چند لایه پنهان و یک لایه خروجی تشکیل شده است. در این ساختار تمام

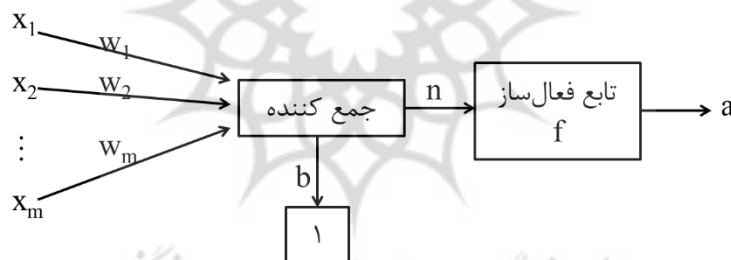
^۱ Multilayer Perceptron

نرون‌های یک لایه به تمام نرون‌های لایه بعد متصلند. شکل ۲ نمایی کلی از یک پرسپترون سه لایه را نشان می‌دهد.



شکل (۲): پرسپترون سه لایه

هر نرون، یک واحد پردازش داده است و نقشی پایه، در کارکرد شبکه عصبی دارد. مدل یک نرون در شکل ۳ آورده شده است. یک نرون مقادیر X_1, X_2, \dots, X_m را به عنوان ورودی دریافت کرده و خروجی a را تولید می‌کند.



شکل (۳): عملیات یک نرون در یک شبکه عصبی

در این مدل، مجموعه‌ای از روابط و اتصالات وابسته به ورودی‌های مختلف X_i مشخص شده است، بطوریکه، هر ورودی X_i دارای وزن w_i می‌باشد. b مقدار بایاس^۱ است که به حاصل جمع ورودی‌های وزن دار اضافه می‌شود (رابطه (۲)). بایاس از تغییرات بزرگ و ناگهانی در خروجی نرون جلوگیری می‌کند و باعث بهبود کارایی شبکه می‌شود. خروجی هر نرون توسط رابطه (۳) محاسبه می‌شود.

$$n = b + w_1 X_1 + w_2 X_2 + \dots + w_m X_m \quad (2)$$

^۱ bias

$$a=f(n) \quad (۳)$$

تابع فعال ساز f ، می تواند خطی یا غیرخطی باشد و بر اساس نیاز مسئله انتخاب می شود. در ادامه توابع فعال ساز به کار رفته در مقاله را معرفی می کنیم.

- تابع خطی: ورودی و خروجی این تابع باهم رابطه مستقیم دارند و همان مقدار ورودی را به عنوان خروجی برمی گردانند.
- تابع سیگموئید^۱: این تابع مقادیر ورودی را در محدوده منفی بی نهایت تا مثبت بی نهایت دریافت کرده و خروجی بین صفر و یک را تولید می کند.

$$a = \frac{1}{1 + \exp(-n)} \quad (۴)$$

- تابع تانزانت سیگموئید^۲: این تابع نیز مشابه تابع سیگموئید است، با این تفاوت که خروجی تولید شده در بازه $(-1, +1)$ قرار می گیرد.

$$a = \frac{2}{(1 + \exp(-2n))} - 1 \quad (۵)$$

آموزش شبکه ها به روش های مختلفی امکان پذیر است که روش پس انتشار (BP) یکی از معروف ترین این روش هاست. در این مقاله جهت آموزش شبکه های موجود، روش پس انتشار با الگوریتم آموزشی لوبنبرگ مارکوارت به کار رفته است، که الگوریتم پس انتشار ساده از روش گرادیان نزولی برای کاهش مجموع مربعات خطا بین خروجی مطلوب و خروجی واقعی، برای همه الگوها استفاده می کند و الگوریتم لوبنبرگ مارکوارت، ترکیبی از روش گرادیان نزولی و روش گاوس نیوتن^۳ را جهت کاهش خطا به کار می بندد.

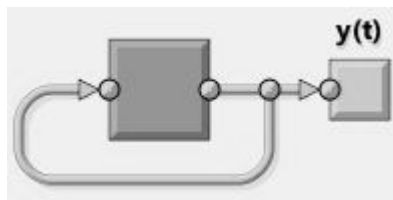
شبکه عصبی NAR

در مقابل روش های آماری جهت پیش بینی سری های زمانی می توان شبکه عصبی مصنوعی NAR که قابلیت پیش بینی فرایندهای اتورگرسیو را داراست، مورد استفاده قرار داد. شبکه عصبی NAR یک شبکه پویا بازگشتی است که دارای چندین لایه است. شکل ۴ نمایی از بازگشتی بودن شبکه عصبی مورد استفاده در این مقاله را نشان می دهد.

¹ Log sigmoid

² Tangent sigmoid

³ Gauss-Newton

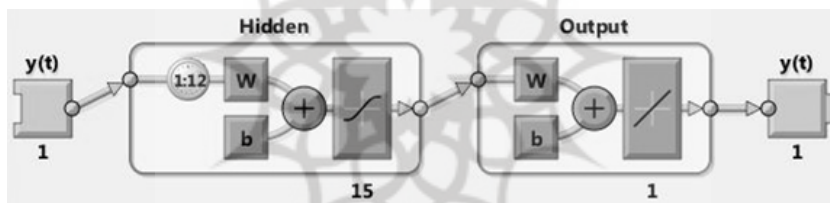


شکل (۴): شبکه عصبی بازگشتی

این شبکه با داشتن مقادیر پیشین سری زمانی از جمله، $x(t-1), x(t-2), \dots, x(t-d)$ قادر به پیش‌بینی مقدار $x(t)$ است.

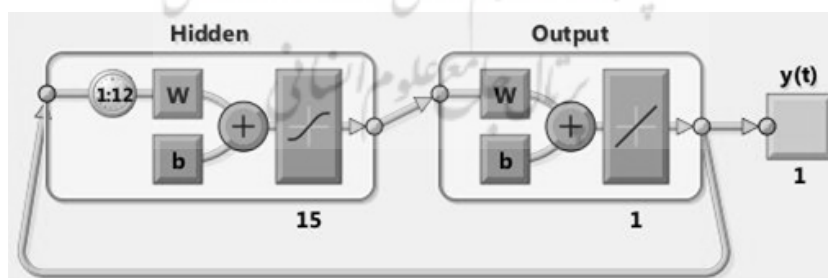
$$x(t) = f(x(t-1), x(t-2), \dots, x(t-d)) \quad (۶)$$

شبکه عصبی مورد استفاده در این مقاله بعد از چندین بار آموزش با تعداد نرون‌ها و تاخیرهای زمانی متفاوت به صورت ۱۵ نرون در لایه میانی و تاخیر زمانی ۱:۱۲ در نظر گرفته شد، چرا که این وضعیت دارای نتایج برتر نسبت به سایر انتخاب‌ها بود. الگوریتم مورد استفاده برای آموزش شبکه روش لونیبرگ مارکوارت (LM) می‌باشد. شکل ۵ معماری شبکه مورد استفاده در این مقاله را نشان می‌دهد.



شکل (۵): معماری شبکه عصبی بکار گرفته شده در این مقاله

بعد از آموزش شبکه، شبکه برای پیش‌بینی مقداری در آینده باید به صورت حلقوی مورد استفاده قرار گیرد. شکل ۶ نشان‌دهنده بستن شبکه به صورت حلقوی است.



شکل (۶): شبکه عصبی بسته

در این روش در هر اجرا با داشتن داده‌های زمان ۱ تا $t-1$ مقدار داده در زمان t را پیش‌بینی می‌کند. در مراحل بعدی با کمک داده‌های پیش‌بینی شده داده‌های جدید پیش‌بینی می‌شوند.

لازم به ذکر است از آنجا که در این روش به کمک داده‌های پیش‌بینی شده، دیگر داده‌ها پیش‌بینی می‌شوند، انتظار می‌رود دنباله به دست آمده همگرا باشد. بنابراین برای پیش‌بینی‌های طولانی مدت، نتیجه‌ای با خطای غیر قابل قبول حاصل خواهد شد. ما براساس بازه‌های ده ساله سرشماری، تا سال ۱۴۰۵ پیش‌بینی را انجام داده که در بخش بعد می‌توان نتایج را مشاهده کرد.

پیش‌بینی تعداد ازدواج و طلاق

برای پیش‌بینی وقایع حیاتی، از ابتدای سال ۱۳۸۶ تا خردادماه ۱۳۹۵، داده‌ها به صورت ماهیانه در دسترس ما قرار داده شد. به کمک شبکه‌های عصبی، تا پایان سال ۱۴۰۴ آمار تعداد ازدواج‌های پیش‌بینی شده در جداول ۲ و ۱ و تعداد طلاق‌های پیش‌بینی شده در جداول ۳ و ۴ قرار داده شده است. در ضمن در شکل‌های ۷ و ۸ نمودار رسم شده بیانگر این آمارها هستند.

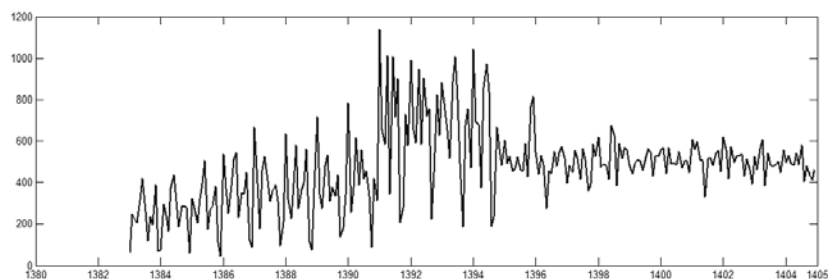


جدول (۱): تعداد ازدواج های ثبت شده استان تا ابتدای سال ۱۳۹۵

ماه/سال	۱۳۸۳	۱۳۸۴	۱۳۸۵	۱۳۸۶	۱۳۸۷	۱۳۸۸	۱۳۸۹	۱۳۹۰	۱۳۹۱	۱۳۹۲	۱۳۹۳	۱۳۹۴
۱	۵۹	۷۷	۳۲۴	۵۳۶	۶۶۶	۶۳۴	۷۱۸	۷۸۲	۱۱۳۸	۹۸۹	۸۸۱	۱۰۴۰
۲	۲۴۷	۲۹۶	۲۷۷	۳۷۴	۳۹۶	۳۱۷	۳۴۴	۲۵۷	۶۵۴	۶۵۴	۷۵۰	۶۹۲
۳	۲۲۳	۲۴۶	۲۰۴	۲۵۲	۱۷۸	۲۲۹	۲۷۶	۳۷۷	۵۸۳	۵۹۱	۶۵۵	۶۸۰
۴	۲۰۴	۱۶۷	۳۲۲	۳۸۶	۴۵۸	۳۹۷	۴۸۳	۶۱۶	۱۰۰۹	۹۴۹	۵۱۶	۳۷۹
۵	۳۱۳	۳۷۸	۳۹۵	۵۱۲	۵۲۵	۵۷۹	۵۳۴	۳۸۹	۳۴۷	۵۸۳	۸۸۷	۸۵۷
۶	۴۱۹	۴۳۷	۵۰۶	۵۴۴	۴۰۹	۲۷۶	۳۱۲	۵۵۸	۱۰۰۸	۹۰۵	۱۰۰۷	۹۷۱
۷	۳۰۲	۳۱۲	۱۷۵	۲۳۳	۳۱۱	۳۶۴	۳۷۸	۴۲۲	۷۱۷	۷۲۳	۸۰۵	۸۴۰
۸	۱۲۱	۱۸۷	۲۶۹	۳۴۹	۳۵۷	۴۰۰	۳۳۶	۴۵۶	۹۰۱	۷۵۵	۴۹۴	۱۸۹
۹	۲۳۸	۲۸۲	۲۸۶	۳۴۶	۳۸۹	۵۶۲	۴۳۷	۳۳۴	۲۱۱	۲۲۳	۱۸۵	۲۳۸
۱۰	۱۹۸	۲۸۸	۳۸۲	۴۴۹	۳۱۲	۱۱۶	۱۴۰	۸۸	۲۸۱	۴۴۷	۶۷۵	۶۶۵
۱۱	۳۸۷	۲۷۸	۱۱۳	۱۲۲	۹۷	۷۸	۱۸۰	۴۲۱	۷۲۸	۸۲۳	۷۵۵	۵۶۴
۱۲	۷۰	۶۱	۴۶	۸۷	۲۱۰	۳۷۵	۳۴۳	۳۱۵	۵۸۲	۶۲۹	۴۷۲	۴۸۳

جدول (۲): پیش بینی تعداد ازدواج های ثبت شده استان تا سال ۱۴۰۵

ماه/سال	۱۳۹۵	۱۳۹۶	۱۳۹۷	۱۳۹۸	۱۳۹۹	۱۴۰۰	۱۴۰۱	۱۴۰۲	۱۴۰۳	۱۴۰۴
۱	۶۰۵	۵۴۸	۳۹۷	۶۱۹	۴۸۸	۵۵۸	۶۰۹	۶۱۹	۵۲۶	۵۰۰
۲	۴۹۶	۴۳۹	۴۸۴	۴۷۴	۴۴۲	۵۶۷	۵۵۹	۵۵۸	۴۶۲	۵۳۱
۳	۵۲۶	۵۲۹	۴۴۸	۴۹۲	۴۹۲	۴۴۵	۵۹۸	۴۲۱	۵۶۴	۴۹۴
۴	۴۵۴	۴۸۶	۵۵۲	۴۷۷	۵۱۱	۵۶۸	۵۱۰	۵۷۲	۶۰۹	۴۸۴
۵	۴۶۷	۲۷۷	۵۱۸	۴۱۶	۵۰۲	۴۸۹	۵۰۷	۴۹۷	۳۸۴	۵۴۱
۶	۵۲۳	۴۵۵	۴۱۶	۶۷۴	۴۶۰	۴۹۶	۳۳۲	۵۲۴	۵۴۱	۴۸۶
۷	۴۶۲	۴۴۲	۵۶۷	۶۲۸	۵۰۷	۴۸۶	۵۱۵	۵۳۰	۴۸۵	۵۸۲
۸	۴۵۷	۵۴۸	۵۱۱	۳۸۵	۵۶۶	۵۵۱	۵۲۰	۵۳۹	۴۷۸	۴۰۵
۹	۵۸۹	۴۷۹	۳۶۲	۵۸۸	۵۴۵	۴۷۲	۴۸۴	۴۳۲	۴۸۷	۴۷۹
۱۰	۴۲۷	۵۴۶	۳۹۳	۵۲۰	۴۳۲	۵۰۲	۵۴۰	۵۱۳	۵۰۳	۴۳۶
۱۱	۷۷۰	۵۷۲	۵۸۴	۵۷۰	۵۲۷	۵۰۸	۵۵۳	۴۶۵	۴۴۹	۴۱۴
۱۲	۸۱۶	۵۱۹	۵۲۵	۵۵۸	۵۲۷	۴۴۸	۴۵۰	۳۹۴	۵۵۵	۴۶۵



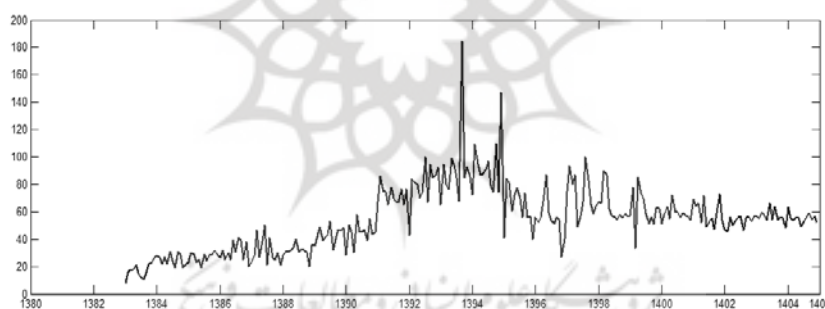
شکل (۷): تعداد ازدواج های ثبت شده استان تا سال ۱۳۹۵ و پیش بینی این تعداد تا سال ۱۴۰۵

جدول (۳): تعداد طلاق های ثبت شده استان تا ابتدای سال ۱۳۹۵

ماه/سال	۱۳۸۳	۱۳۸۴	۱۳۸۵	۱۳۸۶	۱۳۸۷	۱۳۸۸	۱۳۸۹	۱۳۹۰	۱۳۹۱	۱۳۹۲	۱۳۹۳	۱۳۹۴
۱	۸	۲۸	۲۳	۲۶	۲۳	۲۸	۳۵	۲۸	۶۷	۴۳	۶۵	۷۲
۲	۱۶	۲۶	۳۰	۳۲	۲۸	۳۱	۴۲	۵۰	۸۶	۸۴	۹۵	۱۰۹
۳	۱۸	۲۲	۲۹	۲۵	۴۶	۳۱	۴۹	۴۵	۷۵	۸۱	۸۰	۹۵
۴	۱۷	۲۷	۲۳	۳۰	۲۷	۳۲	۳۹	۳۰	۷۵	۸۰	۷۶	۸۷
۵	۲۱	۲۲	۲۵	۲۵	۳۵	۳۴	۴۱	۵۸	۶۵	۷۰	۹۹	۸۸
۶	۱۵	۳۱	۱۹	۳۹	۵۰	۴۰	۴۴	۴۵	۷۸	۷۴	۹۴	۹۱
۷	۱۲	۲۴	۲۹	۳۱	۲۱	۳۱	۵۳	۴۵	۷۰	۱۰۰	۸۵	۹۷
۸	۱۱	۱۹	۲۴	۴۱	۴۱	۳۳	۳۲	۴۷	۶۷	۶۷	۶۷	۷۹
۹	۱۶	۳۱	۲۸	۳۹	۲۶	۳۲	۴۰	۳۹	۶۷	۹۵	۱۸۵	۷۴
۱۰	۲۲	۲۹	۲۹	۲۵	۲۵	۳۰	۴۷	۵۵	۷۷	۸۵	۸۵	۱۱۰
۱۱	۲۳	۱۹	۳۲	۳۸	۳۰	۲۰	۴۶	۴۴	۶۵	۸۷	۹۳	۷۴
۱۲	۲۶	۲۲	۲۹	۲۰	۲۱	۳۵	۴۸	۴۵	۷۷	۹۲	۸۶	۱۴۷

جدول (۴): پیش بینی تعداد طلاق های ثبت شده استان تا سال ۱۴۰۵

ماه/سال	۱۳۹۵	۱۳۹۶	۱۳۹۷	۱۳۹۸	۱۳۹۹	۱۴۰۰	۱۴۰۱	۱۴۰۲	۱۴۰۳	۱۴۰۴
۱	۴۱	۵۶	۷۴	۶۸	۵۷	۵۱	۶۹	۴۷	۵۸	۶۳
۲	۸۵	۵۲	۹۴	۶۷	۷۸	۵۸	۶۴	۴۶	۵۵	۵۵
۳	۸۱	۵۶	۸۰	۹۰	۳۳	۶۴	۶۶	۵۶	۶۰	۵۴
۴	۶۱	۶۸	۸۷	۸۷	۸۵	۵۵	۵۲	۵۰	۵۷	۵۶
۵	۷۲	۸۷	۴۹	۶۲	۷۴	۷۲	۷۱	۵۴	۵۳	۵۶
۶	۷۷	۵۹	۵۷	۵۷	۷۰	۶۰	۴۹	۵۶	۶۶	۴۹
۷	۷۰	۵۴	۶۷	۵۶	۵۸	۶۰	۵۳	۵۷	۵۴	۵۳
۸	۵۶	۵۱	۱۰۰	۵۴	۵۱	۵۶	۵۵	۴۶	۶۴	۵۶
۹	۷۴	۵۶	۸۳	۵۸	۵۶	۵۹	۴۷	۵۵	۵۴	۵۹
۱۰	۵۶	۵۴	۶۵	۵۵	۵۱	۵۸	۶۲	۵۷	۵۶	۵۴
۱۱	۵۷	۲۷	۵۹	۵۹	۶۳	۵۶	۷۳	۵۳	۵۵	۵۶
۱۲	۴۰	۴۰	۶۴	۵۷	۶۲	۵۴	۵۳	۵۷	۴۸	۵۲



شکل (۸): تعداد طلاق های ثبت شده استان تا سال ۱۳۹۵ و پیش بینی این تعداد تا سال ۱۴۰۵

نتیجه گیری

در این مقاله به روش های ریاضی و درون یابی مورد بحث و مزایا و معایب آنها مورد بررسی قرار گرفت. سپس با توجه به مزایای استفاده از شبکه های عصبی، تصمیم بر این شد که این مقاله در بحث پیش بینی از الگوریتم های شبکه های عصبی استفاده کند. شبکه عصبی مصنوعی، که مدل - های ساده شده سیستم نرون های بیولوژیک هستند، از تعداد فراوانی المان های عصبی محاسباتی متصل به هم تشکیل شده اند که قابلیت یادگیری و کسب دانش را دارند و در حل مسائلی مانند مدل سازی سری های زمانی و پیش بینی بر اساس آن ها کاربرد دارند. این شبکه ها قادرند روابط

پیچیده غیرخطی بین متغیرهای ورودی و خروجی را در یک مسئله یاد بگیرند، بدون اینکه دانشی قبلی نسبت به آن مسئله داشته باشند. از این رو در این پژوهش به جای استفاده از مدل‌های ریاضی، شبکه‌های عصبی جهت پیش‌بینی میزان ازدواج و طلاق در سال‌های آتی به کار گرفته شد. از آنجا که تعداد بیشتر داده‌های واقعی، دقت پیش‌بینی را بالا برده و از طرفی دسترسی به اطلاعات واقعی سالهای دور، با توجه به جابه‌جایی‌های مناطق جغرافیایی در تقسیمات کشوری (تقسیم استان خراسان بزرگ به سه استان)، وجود نداشت، داده‌ها در این مقاله به صورت ماهیانه مورد استفاده قرار گرفت. لازم به ذکر است شهرستان طبس در سال ۱۳۹۱ از استان یزد جدا و به استان خراسان جنوبی الحاق شد که با توجه به وسعت شهرستان طبس، خود باعث ایجاد تغییراتی در آمار استان شده است. از آنجا که در پیش‌بینی سریهای زمانی به کمک شبکه‌های عصبی، تمامی داده‌ها مورد استفاده قرار می‌گیرد، این جابه‌جایی خطایی غیر قابل کنترل ایجاد می‌کند. هرچند جابه‌جایی‌های دیگری نیز در استان رخ داده ولی به دلیل این که فقط آمار کل استان مورد بحث قرار گرفته این جابه‌جایی‌ها تأثیری در بحث مورد پژوهش ایجاد نمی‌کند. در نهایت هر چند کار انجام شده نمی‌تواند بدون خطا باشد ولی نسبت به روشهای موجود، بهتر عمل می‌کند.

باید توجه داشت که در بحث ازدواج مناسبت‌های مذهبی در پراکندگی تعداد ازدواجهای ثبت شده در ماههای مختلف سال تأثیر به‌سزایی دارند. به عنوان مثال در ماههای شمسی منطبق بر ماه‌های قمری محرم و صفر ازدواجهای ثبت شده به شدت کاهش و در ماههای شمسی منطبق بر ماه‌های قمری رجب و شعبان ازدواجهای ثبت شده افزایش می‌یابد و این خود یکی از دلایل مهم در ایجاد پرشهای متوالی با ارتفاع زیاد در نمودار ۷ می‌باشد. همانطور که مشاهده می‌شود در نمودار ۸، روند رشد تعداد طلاق‌ها نسبت به ازدواج هموارتر است. در نهایت با وجود تمامی بحثهای فوق انتظار می‌رود روند افزایشی رشد میزان ازدواجهای ثبت شده به نسبت سالهای ۱۳۹۰ تا ۱۳۹۵ حفظ شود و روند افزایشی رشد میزان طلاق به تدریج کاهش یابد.

منابع

- اصغری اسکویی، محمدرضا (۱۳۸۱)، " کاربرد شبکه‌های عصبی در پیش‌بینی سری‌های زمانی"، فصلنامه پژوهش‌های اقتصادی ایران، شماره ۱۲، ۶۹-۹۷.
- حقیقت منفرد، جلال؛ احمد علی نژاد، محمود و متقالچی، سارا (۱۳۹۱)، " مقایسه مدل‌های شبکه عصبی با مدل سری زمانی باکس-جنکینز در پیش‌بینی شاخص کل قیمت سهام بورس اوراق بهادار تهران"، مجله مهندسی مالی و مدیریت اوراق بهادار، شماره ۱۱.
- زمانی، فریبا (۱۳۸۸)، "اهمیت جمعیت‌شناسی در برنامه‌ریزی‌های آمایش سرزمین، با تاکید بر مطالعه مهاجرت‌های استانی"، فصلنامه جمعیت شماره ۷۰/۶۹، ۱۰۷-۱۳۲.
- شکوری، مرتضی (۱۳۸۸)، "آثار تغییرات ساختار نسبی جمعیت ایران بر صندوق بازنشستگی کشوری"، واحد مطالعات و تحقیقات بیمه‌ای موسسه حسابرسی صندوق بازنشستگی کشوری (طرح پژوهشی).
- محمودی، محمدجواد؛ کاظمی‌پور، شهلا؛ احراری، مهدی و نیکونسبتی، علی (۱۳۹۱)، "پیش بینی رشد جمعیت ایران با توجه به مؤلفه‌های اقتصادی-اجتماعی و مبتنی بر رویکرد میان رشته ای"، فصلنامه برنامه‌ریزی و بودجه، سال هفدهم، شماره ۲، ۹۷-۱۲۶.
- میرزائی، محمد و قهفرخی، مهری (۱۳۸۶)، "جمعیت‌شناسی سالمندان در ایران بر اساس سرشماری‌های ۱۳۳۵-۱۳۸۵"، مجله سالمندی ایران، سال دوم، شماره پنجم.
- Diaconescu, E., (2008), "The use of NARX Neural Networks to predict Chaotic Time Series", Wseas Transactions on Computer Research, Issue 3, Volume 3, 182-191.
- Shaarawy S., Khan E., and Elgamel M., (2014), "Modeling and Forecasting the Number of Pilgrims Coming from Outside the Kingdom of Saudi Arabia Using Bayesian and Box-Jenkins Approaches", International Journal of Advanced Computer Science and Applications (IJACSA), Volume 5, No 4, 199-207.