

# مدل تولید سفر بر اساس بکارگیری شبکه‌های عصبی مصنوعی و مقایسه نتایج آن با مدل مبتنی بر رگرسیون

شهریار افندی‌زاده<sup>i</sup>؛ حسین پورخانی بازکیائی<sup>ii</sup>

## چکیده

در این مقاله، موضوع ساخت مدل تولید سفر بر مبنای ساختار شبکه‌های مصنوعی بررسی شده است. در این مقاله ساخت شبکه عصبی پیش‌بینی تولید سفر در قالب شش گام مستقل مورد نظر قرار گرفته و روش استفاده از شبکه عصبی در فرآیند تولید سفر در یک مورد مطالعاتی بکار برده شده است. یکی از مهم‌ترین دست آوردهای این مقاله این است که بکارگیری روش پیشنهادی ارائه شده می‌تواند، طراح مدل را از فرض اولیه در مورد ساختار نهایی رابطه مدل بی‌نیاز سازد. نتایج مدل‌های ساخته شده نشان می‌دهد که راهکار ارائه شده در مقایسه با روش‌های کلاسیک موجود، عملکرد بسیار مناسبی دارد. برای ساخت مدل از آمار و اطلاعات شهر اصفهان استفاده شده است.

## کلمات کلیدی

تولید سفر، مدل‌سازی، شبکه‌های عصبی مصنوعی، پرسپترون چند لایه

## *Trip Generation Modeling Based on Artificial Neural Networks and Comparing With Regression Models*

Shahriar Afandizadeh; Hossein Pourkhani Bazkiayee

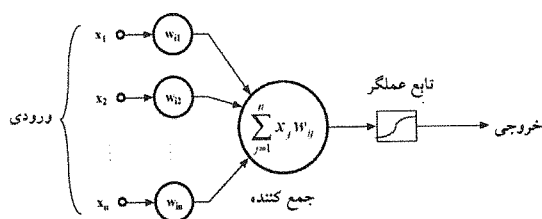
### ABSTRACT

To make a trip generation model based on Artificial Neural Networks is subject of this paper. For applying the model in real network, information of Esfahan city has been used. One of the most important results of this paper is that this method has no need of first assumption of final model relationship at the beginning of modeling. The results has been shown that Artificial Neural Networks is a suitable method compared with classical methods assumption such as regression.

### KEYWORDS

Trip generation, Modeling, Artificial Neural Networks, Multi Layer Perceptron

<sup>i</sup> دانشجوی گروه حمل و نقل دانشکده عمران، دانشگاه علم و صنعت ایران، تلفن: ۰۲۱۷۷۲۴۰۳۹۹ zargari@iust.ac.ir  
<sup>ii</sup> دانشجوی کارشناسی ارشد برنامه ریزی حمل و نقل، دانشگاه علم و صنعت ایران: h\_pourkhani@yahoo.com



شکل (۱): یک مدل با چند ورودی

که در نهایت، خروجی به صورت رابطه (۲) خواهد بود:

$$\text{Out} = f(\text{NET}) \quad (2)$$

توابعی که معمولاً استفاده می‌شوند، توابع خطی، سیگموئیدی، سینوسی، هیپربولیک تانژانت هستند [۴].

یادگیری در مدل‌های انتزاعی شبکه عصبی از طریق قاعده فراگیری پرسپترون امکان پذیر است. منظور از فراگیری دستیابی به بهینه‌ترین وزن‌هاست. قاعده فراگیری پرسپترون چند لایه را قاعده پس انتشار<sup>۲</sup> گویند. مبنای ریاضی الگوریتم پس انتشار، براساس روش بهینه‌سازی کاهش گرادیان<sup>۳</sup> استوار است. گرادیان با علامت مثبت یک تابع، جهتی را در تابع نشان می‌دهد که تابع به طور قابل توجهی افزایش می‌یابد و گرادیان با علامت منفی جهتی را نشان می‌دهد که تابع به طور قابل ملاحظه‌ای کاهش می‌یابد. در الگوریتم انتشار برگشتی، تابع، مقدار خطا و متغیرهای تابع، وزن‌های شبکه‌اند.

در روش بهینه‌سازی یک تابع به صورت رابطه (۳) برای خطا تعریف می‌شود [۱]:

$$E_p = \frac{1}{2} \sum_j (t_{pj} - o_{pj})^2 \quad (3)$$

در این رابطه،  $t_{pj}$  خروجی مدل و  $o_{pj}$  خروجی واقعی است. آنگاه از طریق گرادیان منفی و رابطه (۴)، که قاعده زنجیری نیز بدان اضافه شده است، به یافتن جواب بهینه در فضای اوزان اقدام می‌شود.

$$\frac{\partial E_p}{\partial w_{ij}} = \frac{\partial E_p}{\partial \text{net}_{pj}} \frac{\partial \text{net}_{pj}}{\partial w_{ij}} \quad (4)$$

نتیجه کار برای تابع سیگموئیدی مطابق رابطه (۵) است:

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_{pj} o_{pj} \quad (5)$$

در این رابطه،  $w_{ij}$  نشان‌دهنده وزن بین نرون  $i$ -ام و نرون  $j$ -ام در زمان  $t$  و  $\eta$  نرخ یادگیری<sup>۴</sup> و  $o_{pj}$  خروجی واقعی در گره مورد نظر است. میزان  $\delta_{pj}$  نیز به ترتیب برای لایه‌های میانی و خروجی مطابق با روابط (۶) و (۷) به دست می‌آید (برای تابع سیگموئیدی):

$$\delta_{pj} = o_{pj} (1 - o_{pj}) (t_{pj} - o_{pj}) \quad (6)$$

امروزه مدل‌سازی، بخش مهمی در برنامه‌ریزی حمل و نقل را به خود اختصاص می‌دهد. استفاده از مدل‌های ریاضی در برنامه‌ریزی حمل و نقل باعث توانمند شدن پیش‌بینی‌ها می‌شود. یکی از مدل‌های قوی و کلاسیک در این بخش از مهندسی، مدل‌های چهار مرحله‌ای است که تولید سفر مرحله اول آن می‌باشد. تولید سفر از منظرهای گوناگونی مورد توجه مدل‌سازان قرار گرفته است و راهکارهای گوناگونی برای آن ارائه شده است [۱۰] که مدل‌های الگو برداری شده از طبیعت هم می‌تواند بخش مهمی از آنها باشد. شبکه‌های عصبی مصنوعی مدل‌های الگو برداری شده از سلسله اعصاب انسان هستند که مثال خوبی از الگوهای برگرفته شده از طبیعت است. در این مقاله تلاش می‌شود بین الگوهای پیشنهاد شده برای تولید سفر و سلول‌های عصبی مصنوعی ارتباط برقرار شود و راهکار مناسب برای فرایند تولید سفر مبتنی بر مشخصه‌های اجتماعی اقتصادی جامعه فراهم گردد. در اکثر پروژه‌های مدل‌سازی تولید سفر استفاده از روش رگرسیون از طرف مشاورین مدنظر قرار می‌گیرد. این شیوه در گام نخست مستلزم لحاظ کردن یک ساختار کلی برای مدل و کالیبره کردن آن است. طراح مدل به ناچار برای انطباق داده‌های اولیه در مدل خود مجبور است ساختارهای مختلفی برای ارتباط بین متغیرهای مختلف در نظر بگیرد و سپس آن ساختار را کالیبره کند. این مورد هم زمان بر است و هم به پیشنهاد رابطه‌هایی پیچیده و غیر توجیه پذیر بین متغیرها منجر می‌شود. رابطه (۱۳) نمونه‌ای از این مدل‌هاست. در این مقاله، متدولوژی طراحی شبکه‌های عصبی پرسپترون چند لایه<sup>۱</sup> (MLP) به منظور مدل‌سازی تولید سفر به صورت یک سلسله گام‌های مجزا ارائه شده است.

## ۲- مبانی شبکه‌های عصبی مصنوعی

آموختن در مغز انسان در پی روبرو شدن با رویداد و تعمیم این رویدادها به شرایط جدید است.

در شکل (۱) یک مدل انتزاعی برای نرون با چند ورودی نشان داده شده است. با  $n$  ورودی بردار ورودی  $[x_1, \dots, x_n]$  و وزن‌های سیناپتیکی<sup>۲</sup>  $[w_1, \dots, w_n]$  مجموعه ورودی‌های نرون کامل می‌شود. در ادامه، یک جمله با بردار  $b$  با حاصل‌ضرب ماتریس وزن  $W$  با بردار ورودی  $YX$  جمع می‌شود. ورودی خالص NET به صورت رابطه (۱) حاصل می‌آید:

$$\text{NET}_i = \sum_{i=1}^n X_i W_i + b \quad (1)$$

$$\delta_{pj} = o_{pj} (1 - o_{pj}) \sum_k \delta_{pk} w_{jk} \quad (7)$$

در روش تکمیلی عبارتی به تغییر وزن اضافه می‌شود که متناسب با مقدار تغییر وزن قبلی است. به محض اینکه تنظیم انجام می‌شود، مقدار آن در حافظه ذخیره و برای تعدیل وزن در مرحله بعد به کار گرفته می‌شود. معادلات تنظیم در این روش به صورت روابط (۸) و (۹) است [۹]:

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_{pj} o_{pj} + \alpha \Delta w_{ij}(t) \quad (8)$$

$$\Delta w_{ij}(t) = w_{ij}(t) - w_{ij}(t-1) \quad (9)$$

در معادله (۸)،  $\alpha$  ضریب اندازه حرکت است و مقدار آن در حدود ۰/۹ در تغییر می‌باشد.

در روش بیشترین شیب نزولی، نرخ یادگیری در طول پروسه یادگیری، ثابت و عملکرد الگوریتم به انتخاب درست نرخ یادگیری بسیار حساس است. اگر نرخ یادگیری خیلی بالا باشد شبکه نوسان کرده و آموزش نمی‌بیند. اگر نرخ یادگیری خیلی پایین باشد الگوریتم شبکه بسیار کند به جواب نهایی همگرا می‌شود [۵]. در حقیقت، روش نرخ یادگیری متغیر سعی می‌کند نرخ یادگیری را تا حد ممکن بالا نگه دارد. در هر گام از روند یادگیری ابتدا شبکه با نرخ یادگیری موجود آموزش می‌بیند و خروجی‌های شبکه و خطای شبکه محاسبه می‌شود اگر خطای epoch فعلی نسبت به خطای epoch قبلی از یک نسبت مشخصی؛ که به وسیله طراح پیشنهاد می‌شود، بیشتر باشد، وزن‌ها و بایاس‌های موجود به همراه نرخ یادگیری؛ قبول نمی‌شوند و در یک ضریب کاهشدهنده؛ که به وسیله طراح پیشنهاد می‌شود، ضرب می‌گردند. اگر خطای فعلی نسبت به خطای قبلی کمتر باشد، نرخ یادگیری با یک ضریب افزایشی؛ که توسط طراح پیشنهاد می‌شود، افزایش می‌یابد.

در روش برجسته فقط علامت مشتق برای تعیین جهت به‌روز رسانی وزن‌ها استفاده می‌شود و اندازه مشتقات تأثیری در به‌روز رسانی وزن‌ها ندارد. مقادیر به‌روز رسانی برای وزن‌ها و بایاس‌ها به وسیله یک فاکتور عمل‌کننده تحت تأثیر قرار می‌گیرد. اگر مشتق تابع عمل‌کننده نسبت به وزن‌ها دارای علامت مشابه برای دو تکرار متوالی باشد؛ این فاکتور، افزایشی، در غیر این صورت، فاکتور کاهشدهنده خواهد بود. در صورتی که مشتق صفر باشد به روز رسانی بی‌تغییر می‌ماند.

کلیه الگوریتم‌های پس‌انتشار پایه، وزن‌ها را در مسیری با بیشترین شیب نزولی تغییر می‌دهد که این، مسیری است که در آن تابع عمل‌کننده با سرعت بیشتری کاهش می‌یابد. این روش الزاماً موجب همگرایی سریع نمی‌شود. در روش گرادینان مزدوج<sup>۱۱</sup> جست و جوی مسیر بهینه در طول جهت‌های در هم

آمیخته صورت می‌گیرد که عموماً همگرایی بیشتری را نسبت به روش بیشترین شیب نزولی در پی دارد. روش‌های گرادینان مزدوج با جست و جو در تندترین شیب نزولی شروع می‌شوند. پروسه عمومی برای تعیین مسیر جست و جوی جدید برای ترکیب با مسیر تندترین شیب نزولی با هماهنگی با مسیرهای جست‌وجوی قبلی به صورت رابطه (۱۰) خواهد بود [۴].

$$p_k = -g_k + \beta_k p_{k-1} \quad (10)$$

در این رابطه،  $P_k$  شیب جست‌وجو در مرحله  $k$ ،  $g_k$  بیشترین شیب در مرحله  $k$ ؛  $\beta_k$  نسبت مربع نرم‌گرادینان فعلی به مربع نرم‌گرادینان قبلی است.

گونه دیگری از روش‌های گرادینان مزدوج برای بهینه‌سازی سریع، وجود دارد که عموماً در این روش تغییر وزن از رابطه (۱۱) پیروی می‌کند [۹]:

$$p_{k+1} = p_k - A_k^{-1} g_k \quad (11)$$

این رابطه،  $A_k$  ماتریس هسیان<sup>۱۲</sup> تابع عملکردی نسبت به مقادیر جاری وزن‌ها و بایاس‌هاست. البته، محاسبه ماتریس هسیان برای شبکه‌های عصبی بسیار مشکل و پیچیده است. روش‌های مختلفی بر مبنای روش نیوتن وجود دارد که به محاسبه مشتق دوم نیازی ندارند و روش گوس - نیوتن<sup>۱۳</sup> نامیده می‌شوند. این روش‌ها تقریبی از ماتریس هسیان را در هر دوره تکرار از الگوریتم بکار می‌برند.

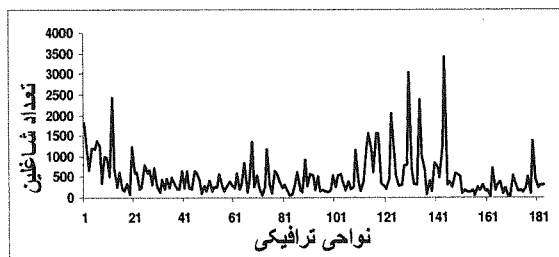
الگوریتم دیگری که در این قسمت به آن اشاره می‌شود، روش لوبنبرگ-مارکوآرت<sup>۱۴</sup> است. روش الگوریتم لوبنبرگ - مارکوآرت به جای استفاده از ماتریس هسیان، از ژاکوبین<sup>۱۵</sup> استفاده می‌کند. این کار از طریق رابطه (۱۲) صورت می‌گیرد [۹]:

$$P_{k+1} = P_k - [J'J + \mu I]^{-1} J'e \quad (12)$$

در این رابطه،  $J$  ماتریس ژاکوبین است که شامل مشتق اول از خطاهای شبکه نسبت به وزن‌ها و بایاس‌هاست.  $J'$  ترانزپانده ماتریس ژاکوبین،  $e$  بردار خطای شبکه است.  $\mu$  یک پارامتر عملگر می‌باشد. اگر  $\mu$  بالا باشد کاهش گرادینان با گام‌های کوچک خواهد بود.  $I$  ماتریس یکانی است. از آنجا که روش نیوتن روش سریعی می‌باشد و در نزدیکی خطای مینیمم صحیح‌تر عمل می‌کند؛ لذا هدف این است که هرچه سریع‌تر به سمت روش نیوتن سوق پیدا کنیم. لذا  $\mu$  پس از هر گام موفق کاهش می‌یابد (کاهش در تابع عملگر) و در صورتی که گام اولیه تابع عملگر را افزایش دهد، پارامتر  $\mu$  افزایش می‌یابد. در این روش، تابع عملگر در هر تکرار از الگوریتم کاهش می‌یابد.

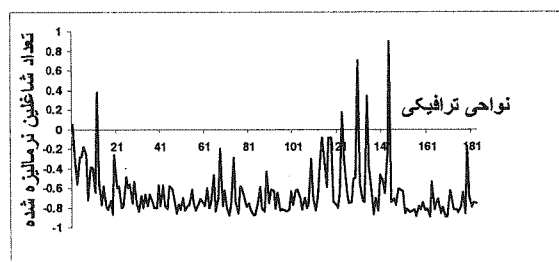
شایان ذکر است در تعبیری از شبکه‌های عصبی از این شبکه‌ها به عنوان "جعبه سیاه" نام برده می‌شود [۸]. شکل (۲) به طور شماتیک این امر را نشان می‌دهد. از آنجا که در پروسه

داده‌ها را ببینیم.



شکل (۳): توزیع داده‌های تعداد شاغلین نواحی ترافیکی

شکل (۳) به داده‌های برگرفته شده از بانک اطلاعاتی (به طور مستقیم) مربوط است. همین داده‌ها اگر به بازه [۱-۱۰۱] نگاشت شوند توزیعی مانند شکل (۴) پیدا می‌کنند. همان طور که مشاهده می‌شود توزیع نسبی بین داده‌ها تغییری نکرده و فقط بازه عملکردی آنها تغییر کرده است.



شکل (۴): توزیع داده‌های تعداد شاغلین نواحی ترافیکی به صورت نگاشت شده در بازه [۱-۱۰۱]

در مورد مدل‌هایی که براساس تابع سیگموئید بکار می‌روند معمولاً بازه [۱ و ۰] و برای مدل‌هایی که از تابع تانژانت هیپربولیک استفاده می‌کنند بازه [۱ و -۱] مدنظر قرار می‌گیرد. از آنجا که مدل ساخته شده به احتمال زیاد باید بعدها برای داده‌هایی خارج از محدوده‌ای که در حال حاضر متغیرها در آن قرار دارند، روبرو شود؛ لذا پیشنهاد شده است که محدوده‌هایی از بازه نگاشت (چپ و راست بازه) که احتمال روبرو شدن با داده‌هایی فراتر از آن محدوده وجود دارد کمی کوچک تر در نظر گرفته شوند. این مورد در شاخه‌های دیگر مهندسی نیز بخوبی جواب داده است؛ لذا برای این قسمت بازه [۰/۹، -۰/۹] پیشنهاد می‌شود.

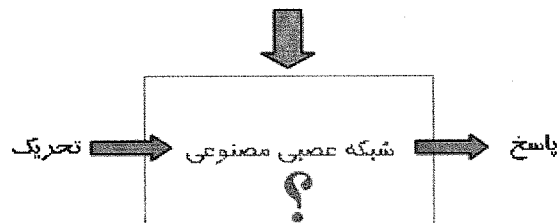
### ۳-۳- چگونگی دسته‌بندی داده‌ها به دسته‌های

#### آموزش و آزمون

لازم است شبکه آموزش دیده، از طریق داده‌هایی که به پروسه مدل‌سازی وارد نشده‌اند ارزیابی شود. معمولاً جمعیتی بین ۱۰ تا ۲۰ درصد از کل داده‌ها برای آزمون شبکه اختصاص داده می‌شود. انتخاب کاملاً تصادفی داده‌های آزمون ممکن است باعث شود داده‌های خیلی تأثیرگذار از داده‌های آموزش

مدل‌سازی، شبکه‌های عصبی برای فرم کلی توابع هدف هیچ پیش‌فرضی قائل نیستند و از آنجا که تعداد لایه‌های پنهان و تعداد نرون‌های بکار رفته در هر لایه به عهده طراح است؛ لذا امکان ایجاد انواع مدل‌ها با سطوح عملکردی متفاوت و همگرایی مختلف به جواب فراهم است.

تنظیم کننده‌ها



شکل (۲): نمایی شماتیک از جعبه سیاه

### ۳- فرایند تحقیق

در این قسمت سعی می‌شود روش‌هایی برای انتخاب بهینه عناصر سازنده شبکه عصبی ارائه شود.

#### ۳-۱- انتخاب تابع یادگیری مناسب

تابع هیپربولیکی در مقایسه با تابع سیگموئیدی معمولاً نتایج مطلوب‌تری را به دنبال دارد. این مورد شاید به این دلیل باشد که در صورت استفاده از تابع تحریک تانژانت هیپربولیک بازه دلخواه برای نگاشت مقادیر گسترده‌تر از حالتی است که تابع سیگموئید بکار برده می‌شود و ورودی‌ها برای شبکه قابل تمایزتر هستند. از نگاهی دیگر، بیولوژیست‌ها از تابع تحریک تانژانت هیپربولیک به عنوان مدل ریاضی از تحریک سلول عصبی استفاده می‌کنند [۲] این مورد این نکته را به ذهن می‌رساند که احتمالاً بکارگیری این نوع تابع در شبکه عصبی خصوصیتی را به نمایش می‌گذارد که هماهنگ با سیستم بیولوژیکی خواهد بود.

#### ۳-۲- نگاشت زوج‌های آموزشی به بازه‌های مناسب

برای خیلی از موارد عملی، انتخاب واحدهای اندازه‌گیری مورد استفاده برای هر یک از متغیرهای ورودی می‌تواند به این منجر شود که محدوده تغییر داده‌ها تفاوت‌های قابل توجهی در مورد هر یک از متغیرهای بکار رفته در مدل داشته باشد. بکارگیری مستقیم چنین داده‌هایی به طور مستقیم در شبکه عصبی ممکن است موجب بروز حساسیت به اعداد بزرگ تر (از نظر عددی) شود [۷]. این مشکل می‌تواند از طریق نگاشت زوج‌های آموزشی به محدوده‌های مطلوب، از حالت بحرانی خارج شود. در شکل (۳) و (۴) می‌توانیم نمونه‌ای از نگاشت

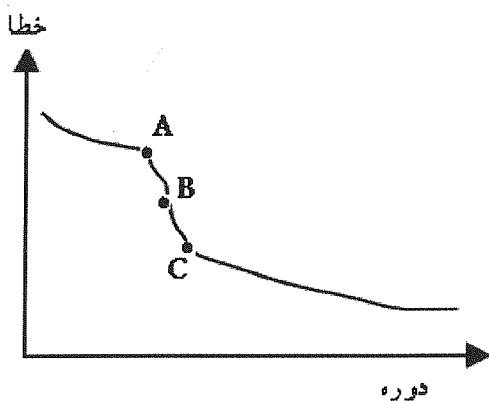
حذف شوند.

[0/5, 0/5]، [0/7, 0/7] و [0/9, 0/9] ساخته شد. در هر مورد به تعداد ۳۰ مورد وزن‌دهی‌ها صورت گرفت تا ماهیت وزن‌دهی تصادفی موجب گرایش به نتایج نادرست نشود. در نهایت، بازه [0/5, 0/5] مناسب تشخیص داده شد.

### ۳-۶- تصمیم‌گیری در مورد انتخاب الگوریتم آموزش

#### شبکه

شاید مهم‌ترین و حساس‌ترین تصمیم‌گیری در پروسه طراحی شبکه‌های عصبی تصمیم‌گیری در مورد الگوریتم آموزش شبکه باشد. به طور کلی، روند آموزشی به صورت شماتیک به صورت شکل (۵) است [۶]:



شکل (۵): روند کلی آموزش شبکه‌های عصبی

شبکه، یک مناسب سازی اولیه برای اوزان خود خواهد داشت (قبل از نقطه A): سپس فرآیند یادگیری انجام می‌شود (از نقطه A تا C) و در نهایت، شبکه به بهینه سازی نهایی برای اوزان اقدام می‌کند.

در این تحقیق مشاهده شد که اگر شیب نمودار آموزشی خیلی تند باشد این امر موجب آن می‌شود که محدوده تصمیم‌گیری در مورد ساختار بهینه شبکه کوچک باشد. آنچه که برای سازنده مدل مهم است صرفاً دستیابی به شبکه‌ای نیست که به کمترین خطای ممکن همگرا شده باشد. اگر دو الگوریتم آموزش مختلف با هم مقایسه شوند، الزاماً نمی‌توان نتیجه گرفت، شبکه‌ای که به جواب کمتری همگرا شده شبکه مطلوب‌تری است؛ چرا که ممکن است چنین شبکه‌ای روابط بین متغیرها را از بر کرده باشد نه اینکه آموخته باشد.

بررسی‌های مدل‌های ساخته شده نشان داد روش الگوریتم برجهنده از آنجا که بهترین رفتار همگرایی به جواب نهایی را دارد؛ می‌تواند بهترین انتخاب باشد. الگوریتم‌های مبتنی بر "بیشترین شیب نزولی" و "بیشترین شیب نزولی با مومنتم" به خطای مورد قبولی همگرا نمی‌شوند و جواب چندان قابل قبولی ارائه نمی‌دهند. الگوریتم مارکوارت - لونیبرگ نیز هر چند به خطای خیلی کمی همگرا می‌شود؛ ولی از آنجا که کل پروسه یادگیری آن در چند دوره اول اتفاق می‌افتد بستر مناسبی برای

روش پیشنهادی برای این موضوع این است که ابتدا شبکه به کمک کلیه داده‌ها آموزش ببیند سپس به هر یک از داده‌ها از نظر انطباق بین اندازه هدف و اندازه محاسبه شده به وسیله مدل، امتیاز نسبی داده شود؛ سپس داده‌ها براساس امتیازی که کسب کرده‌اند رتبه‌بندی شوند؛ آنگاه به چند دسته مساوی تقسیم، و از میان هر یک از دسته‌ها به تصادف ۲۰ درصد داده‌ها جدا، و برای دفعات بعدی تحلیل شبکه به عنوان داده‌های آزمون در نظر گرفته شوند. اینکار، این احتمال را که داده‌های گوناگون با درجات مختلفی از هماهنگی با کل مجموعه، شانس مساوی برای وارد شدن در پروسه تحلیل و یا قرار گیری در دسته آزمون را داشته باشند، فراهم می‌کند.

### ۳-۴- ساختار شبکه

منظور از ساختار شبکه عصبی مصنوعی، تعداد لایه‌های آن و همین‌طور تعداد نرون‌ها در هر یک از لایه‌ها است. هیچ توجیه ریاضی برای تعداد لایه‌های مناسب برای ساخت یک شبکه عصبی مصنوعی و همین‌طور تعداد نرون‌ها وجود ندارد. این مورد از سوی برخی محققین، یکی از نقاط ضعف شبکه‌های عصبی محسوب می‌شود [۵]. با یک روش مناسب می‌توان در مورد تعداد لایه‌های مخفی به درستی اظهار نظر کرد. با بررسی‌های انجام شده در این تحقیق و با ساخت مدل‌های مختلف و ارزیابی آنها مشاهده شد که افزودن به تعداد لایه‌های شبکه عصبی از قدرت تعمیم دهی شبکه می‌کاهد. روش کار بدین صورت است که شبکه‌هایی با یک، دو و سه لایه پنهان با تلفیقی از تعداد نرون‌های مختلف در آنها ساخته و ارزیابی شد که از بررسی عملکرد شبکه‌های مذکور این نتایج به دست آمد: برای مدل‌سازی مدل‌های تولید سفر، شبکه‌های تک لایه در اولویت هستند و شبکه‌های دو لایه نیز در صورت لزوم مد نظر قرار می‌گیرند. در نظر گرفتن شبکه‌های با سه لایه مخفی و یا بیشتر چندان مطلوب نیست.

### ۳-۵- وزن دهی اولیه به شبکه

ارائه یک مدل شبکه با ساختار مناسب، با دستیابی به مقادیر مناسب برای وزن‌های بین لایه‌ها و بایاس‌های نرون‌ها تحقق می‌یابد. روش‌های بکارگیری برای رسیدن به جواب بهینه در گام نخست به داده‌های اولیه مبتنی هستند که اصولاً کاملاً تصادفی اختیار می‌شوند. اگر بخواهیم در مورد یک شبکه عصبی، دو مجموعه ماتریس مختلف را مدنظر قرار دهیم، الزاماً این‌گونه نخواهد بود که هرکدام از دسته ماتریس‌ها؛ که بعد از اعمال بر شبکه جواب بهتری را به دست بدهند، بعد از سیکل‌های آموزش نیز جواب بهتری را به ما بدهند (این مورد با ساخت شبکه عصبی متنوع و آموزش آنها با دسته‌های مختلف اوزان نتیجه گرفته شد). بدین‌منظور، شبکه‌های مختلف تولید سفر با وزن‌دهی اولیه در بازه‌های [0/1, 0/1]، [0/3, 0/3]،

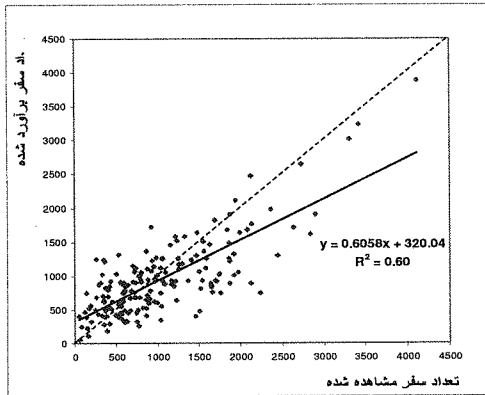
طراحی فراهم نمی‌کند و طراح فرصت کافی برای انتخاب‌های بهینه ندارد.

#### ع- مطالعه موردی

در این تحقیق، شهر اصفهان به عنوان مطالعه موردی انتخاب شده است. بانک اطلاعاتی مورد استفاده در این تحقیق بر مبنای مطالعات مبدأ، مقصد سال ۱۳۷۹ بوده است. علاوه بر بانک مفید مذکور، به بانک اطلاعاتی شهر نیاز است که جدا از بحث مبدأ، مقصد سفرها، خصوصیات اجتماعی اقتصادی ناحیه‌های ترافیکی شهر را بخوبی پوشش دهد. منبع این بانک اطلاعاتی مرکز آمار و اطلاعات ایران بوده است. دلیل انتخاب شهر اصفهان، در دسترس بودن آمار و اطلاعات مطالعات مبدأ مقصد این شهر بوده است. از آنجا که مدل ارائه شده برای جذب سفرهای تفریحی؛ که از طرف مشاور وقت ارائه شده بوده، از پیچیدگی خاصی برخوردار بود و شاخص  $R^2$  به دست آمده برای آن نسبتاً کم است؛ این سفرها برای مدل‌سازی انتخاب شده است. مدل پیشنهادی مبتنی بر رگرسیون برای جذب سفرهای تفریحی؛ که مشاور وقت پیشنهاد کرده است، به صورت رابطه (۱۳) است [۳]. شاخص  $R^2$  برای این مدل برابر ۰/۶ بوده است.

$$D_j^{tr} = 0.041 p_j + 37/917 APARK_j + 11/564(\text{Min}(vkj, 800)).(CAR_j) + (1101 dt_j + 1288 d65j + 766 d_{106j}).p \quad (13)$$

در این رابطه  $D_j^{tr}$ ، میزان سفرهای جذب شده به ناحیه  $J$ ، جمعیت ناحیه  $J$ ،  $CAR_j$  سرانه مالکیت خودرو در ناحیه  $J$ ،  $VK_j$  تعداد واحدهای کسبی در ناحیه  $J$ ،  $APARK_j$  مساحت پارک‌های تفریحی ناحیه ترافیکی  $J$  (هکتار)،  $dt_j$  برابر ۱ اگر  $J \in \{12, 27, 113\}$  و غیر از آن برابر صفر؛  $d65j$  برابر ۱ اگر  $J = 65$  و در غیر این صورت برابر صفر؛  $d106j$  برابر ۱ اگر  $J = 106$  و در غیر این صورت برابر صفر. نمودار برآزش مدل پیشنهادی برای مقادیر مشاهده شده و خروجی مدل در شکل (۶) نشان داده شده است.



شکل (۶): نمودار برآزش مدل رگرسیون برای جذب سفر تفریحی مبتنی بر نواحی

با توجه به مطالب بخش ۳ این مقاله، در این مطالعه موردی، شبکه‌ای با یک لایه مخفی شامل ۷ نرون برای مدل‌سازی جذب سفرهای تفریحی مبتنی بر نواحی ترافیکی طراحی شد. لایه ورودی شامل ۴ نرون معادل ۴ متغیر  $P_j$ ،  $CAR_j$ ،  $VK_j$  و  $APARK_j$  است. شاخص  $R^2$  برای داده‌های آموزش و آزمون به ترتیب برابر ۰/۶۸ و ۰/۶۳ به دست آمد (البته این مدل‌ها، مدل‌های اولیه هستند و مدل‌هایی که در ادامه به آنها اشاره می‌شود از توانایی‌های بیشتری برای درون‌یابی برخوردارند) که نشان‌دهنده قدرت رقابت شبکه‌های عصبی در مقابل مدل‌های رگرسیون می‌باشد. نکته قابل توجهی که در مورد مدل سفرهای تفریحی جذب شده به ناحیه‌ها شاخص وجود دارد، وجود متغیرهای کمکی<sup>۱۵</sup> داخل بدنه مدل است. وجود چنین متغیرهایی مدل را نسبت به حدود داده‌های یک ناحیه کاملاً حساس می‌کند. این مورد، این امر را به ذهن متبادر می‌سازد که شاید بهتر باشد مدل شبکه عصبی تولید سفر (جذب سفر) براساس کل داده‌ها ساخته شود. ارزیابی مدل در این حالت، منوط به مطالعات سال‌های بعد است. در تحقیق حاضر، پراکندگی داده‌ها برای چهار پارامتر اصلی تأثیرگذار در مدل چنان است که حذف هر یک از رکوردها از دسته آموزش و تخصیص آن به دسته آزمون موجب بروز خطای زیاد در مدل‌سازی می‌شود. به منظور توضیح بیشتر، از یک نوع رتبه‌بندی داده‌ها استفاده می‌شود که در طی آن داده‌ها به ۴ رتبه تقسیم می‌شوند. نخست در مورد هر پارامتر از داده‌ها میانگین‌گیری می‌شود، داده‌های بالای میانگین و پایین میانگین تشکیل دو دسته از داده‌ها را می‌دهند؛ سپس در هر یک از دسته‌ها میانگین‌گیری صورت می‌گیرد و باز هم میانگین هر دسته ملاک تقسیم‌بندی قرار می‌گیرد؛ بنابراین داده‌ها به ۴ دسته تقسیم می‌شوند. در گام بعدی به داده‌های مربوط به هر پارامتر از کم تا زیاد از A تا D رتبه داده می‌شود. رتبه مربوط به هر رکورد (ناحیه ترافیکی) تلفیقی از رتبه‌های کسب شده از طرف ۴ پارامتر اصلی مربوط به همان رکورد می‌باشد. برای

داده‌ها، تنها نماینده رتبه خود در داده‌ها هستند؛ لذا می‌بایست رویکرد ما به مدل‌سازی برای این نوع داده‌ها رویکردی مبتنی بر بکارگیری کلیه داده‌ها باشد تا هر یک از داده‌ها این فرصت را داشته باشند تأثیر خاص خود را بر مدل بگذارند.

اینکه مساله مورد نظر با انواع داده‌ها از نظر پراکندگی روبرو شود لازم است داده‌های مربوط به کلیه رکوردها، وارد مدل شوند و اختصاص برخی از داده‌ها در دسته آزمون؛ که به قیمت حذف آنها از دسته آموزش می‌شود، ممکن است خطای زیادی در مدل‌سازی ایجاد کند. مخصوصاً اینکه ۲۴ درصد از

جدول (۱): رتبه‌های مختلف دیده شده در داده‌های ورودی مدل جذب سفر تفریحی

رتبه	تعداد	رتبه	تعداد	رتبه	تعداد	رتبه	تعداد	رتبه	تعداد	رتبه	تعداد
AADA	۱	BDBC	۱	DACC	۱	AAAA	۲	BDAA	۳	CABA	۵
ABBC	۱	CABB	۱	DACD	۱	AABA	۲	CAAA	۲	CBBA	۵
ACAC	۱	CADA	۱	DBAA	۱	ABBA	۲	CACA	۲	DABA	۵
ACBA	۱	CBAC	۱	DBAC	۱	ACCA	۲	CBAA	۳	BAAA	۵
ACDC	۱	CBBC	۱	DBBC	۱	ADBC	۲	CCAA	۳	CCBA	۶
ACDD	۱	CBCC	۱	DBCBC	۱	BABA	۴	CDBA	۳	DBBA	۶
ADAC	۱	CBCD	۱	DCAA	۱	BBDA	۲	ACAA	۴	BBCA	۷
ADAD	۱	CBDC	۱	DCAD	۱	BDAC	۲	ACDA	۴	BCAA	۷
ADDA	۱	CCAC	۱	DCBA	۱	CCAD	۲	BDCA	۴	DACA	۷
BACA	۱	CCBB	۱	DCBB	۱	DAAA	۲	CBCA	۴		
BADA	۱	CDBD	۱	DCBD	۱	DCBC	۲	CCCA	۴		
BBBC	۱	CDCD	۱	DCCA	۱	ABCA	۳	ABAA	۵		
BCAC	۱	DAAB	۱	DCCB	۱	BBAA	۳	ADAA	۵		
BCBA	۱	DABC	۱	DDCA	۱	BCCA	۳	BBBA	۵		

داده‌های جدید دارد. با توجه به این نکات، به ساخت شبکه‌های مختلف با سطوح عملکردی متفاوت برای مدل‌سازی میزان جذب سفر تفریحی اقدام شد که در جدول (۲) مشخصات شبکه‌ها و شاخص  $R^2$  برای آنها ارائه شده است. در این مرحله، همه شبکه‌های ساخته شده با یک لایه مخفی بوده‌اند و از الگوریتم مارکوآرت- لوینبورگ استفاده شد؛ چرا که این الگوریتم بهترین همگرایی به جواب نهایی را دارد و از طرفی رویکرد ما در این قسمت بر ارزیابی از طریق دسته آزمون مبتنی نیست؛ بلکه کل داده‌ها وارد پروسه مدل‌سازی می‌شوند. جدول (۲) نتایج مربوط به ساختارهای مختلف شبکه‌های عصبی به منظور دستیابی به شبکه‌هایی با سطوح عملکردی مختلف را در بر دارد.

همان گونه که در بخش مبانی شبکه‌های عصبی توضیح داده شد در تعبیری از شبکه‌های عصبی از این شبکه‌ها به عنوان "جعبه سیاه" نام برده می‌شود. با این رویکرد، از آنجا که طراح شبکه امکان ساخت شبکه‌های مختلف با سطوح عملکردی متفاوت از نظر درونیابی در فضای جواب‌ها را دارد؛ لذا می‌تواند به جای ارائه یک ساختار شبکه و ارزیابی آن با داده‌های موجود؛ به ارائه چندین شبکه با ساختارهای مختلف اقدام کند. در این مرحله، طراح با بکارگیری عناصر سازنده مختلف، شبکه‌هایی با توانایی‌های مختلف در درونیابی ایجاد می‌کند؛ چرا که طراح اگر قرار باشد فقط به ساخت شبکه‌ای با بهترین همگرایی به جواب (با کل داده‌ها) بسنده کند این کار ممکن است باعث شود شبکه مذکور توان درونیابی در فضای جواب‌ها را نداشته و فقط روابط بین داده‌ها را حفظ کرده باشد؛ اما ارائه چند شبکه با توانایی‌های مختلف، در همگرایی به خطای کمتر، این اطمینان را ایجاد می‌کند که در صورت عدم قبول عملکرد یک شبکه، شبکه‌ای با ساختاری ساده‌تر وجود دارد که هرچند به خطایی بزرگ‌تر همگرا می‌شود؛ ولی توانایی‌های بهتری در درونیابی در فضای جواب‌ها برای

جدول (۲): شبکه‌های عصبی مختلف با سطوح عملکردی متفاوت.

شاخص $R^2$	ساختار پیشنهادی
۰/۶۴	۴-۵-۱
۰/۷	۴-۷-۱
۰/۷۵	۴-۹-۱
۰/۷۹	۴-۱۱-۱
۰/۸۲	۴-۱۳-۱
۰/۸۶	۴-۱۵-۱
۰/۹	۴-۱۷-۱
۰/۹۲	۴-۱۹-۱
۰/۹۳	۴-۲۱-۱

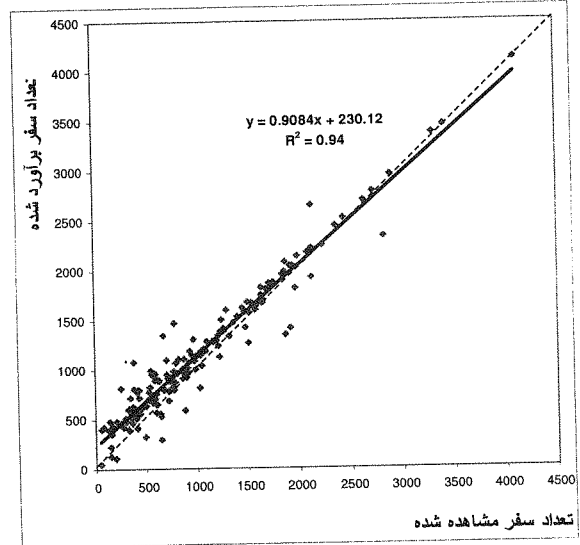
$R^2$  برای آن بیش از ۰/۹ به دست آمد در حالی که در بهترین حالت در روش رگرسیون برابر ۰/۶ بوده است.

۴- بحث در مورد دیدگاه ویژه‌ای برای شبکه‌های عصبی، که شبکه عصبی را جعبه سیاه نام نهاده و استفاده از این دیدگاه به منظور مدل‌سازی شبکه‌های مختلف با سطوح متفاوت عملکردی برای افزایش انعطاف‌پذیری با داده‌های آینده.

۵- نشان دادن این مطلب که شبکه‌های عصبی چنان از قدرت درونی‌یابی توانمندی برخوردارند که در پروسه مدل‌سازی بدون نیاز به متغیرهای کمکی نتایج کاملاً قابل قبولی ارائه می‌دهند. (برعکس مدل‌های رگرسیون که گاهی اوقات از متغیرهای کمکی استفاده می‌کنند.)

۶- معرفی شبکه‌های عصبی به عنوان راهکاری که برای ایجاد ارتباط مابین چند متغیر به فرض کردن هیچ ساختاری برای مدل نهایی نیاز ندارد؛ و بحث کالیبره کردن مدل که در روش‌های کلاسیک دیده می‌شود تبدیل به آموزش شبکه می‌شود و امکان دستیابی به هر ساختار پیچیده‌ای از طریق شبکه‌های عصبی فراهم می‌شود.

برای نمونه، نمودار برآزش شبکه با ساختار ۴-۲۱-۱ در شکل (۷) نشان داده شده است. مقایسه اشکال (۶) و (۷) بخوبی توانایی‌های شبکه عصبی را در مدل‌سازی جذب سفرهای تفریحی نشان می‌دهد. ساختار ۴-۲۱-۱ به معنی شبکه‌ای با ۴ نرون در لایه ورودی، ۲۱ نرون در لایه پنهان و ۱ نرون در لایه خروجی است.



شکل (۷): نمودار برآزش مدل شبکه عصبی طراحی شده برای جذب سفرهای تفریحی مبتنی بر نواحی

## ۶- مراجع

- [۱] البرزی، محمود، "آشنایی با شبکه‌های عصبی"، موسسه انتشارات علمی دانشگاه علمی دانشگاه صنعتی شریف، ۱۳۸۰
- [۲] غضنفری، محمود، شبکه‌های عصبی (اصول و کاربردها)، دانشگاه علم و صنعت ایران، تهران، ۱۳۸۲.
- [۳] گزارش فاز یک مطالعات جامع حمل و نقل شهری اصفهان (گزارش نهایی نتایج آمارگیری مبدأ، مقصد)، سازمان حمل و نقل و ترافیک اصفهان ۱۳۷۹.
- [۴] مناج، محمدباقر؛ "هوش محاسباتی (جلد اول: مبانی شبکه‌های عصبی)"، انتشارات دانشگاه صنعتی امیرکبیر، تهران، ویرایش اول، ۱۳۷۹.
- [۵] Abhijit Dharia, Hojjat Adeli " Neural network model for rapid forecasting of freeway link travel time" Engineering Applications of Artificial Intelligence 16 (2003) 607-613.
- [۶] Fausset ,L, "Fundamental Of Neural Networks Architectures Algorithms and applications", Prentice Hall, 1994.
- [۷] Hong Xing, Li , "Fuzzy Neural Intelligent Systems" , Han Pang Huang , CRC.Press llc , 2001
- [۸] H. Murat Celik , " Modeling freight distribution using artificial neural networks" Journal of Transport Geography 12 (2004) 141-148.
- [۹] Matan ,Gupta ; Liang , Jin , Foreword by Lotfi.Zadeh , "Static And Dynamic Neural Networks" , A John Wily & Sons , Inc 2003.
- [۱۰] Ourtuzar , Juan de Dios , " Modelling Transport 3rd Edition " , John Wily & Sons Ltd 2001.

## ۵- نتیجه

بطور کلی، ره‌آورد این تحقیق را می‌توان در قالب چند بخش برشمرد:

- ۱- ارائه راهکارهایی در زمینه انتخاب عناصر سازنده شبکه عصبی
- ۲- ساخت شبکه عصبی برای مدل‌سازی جذب سفر تفریحی
- ۳- مقایسه مدل‌های شبکه عصبی با مدل‌های رگرسیون و نتیجه‌گیری این مطلب، که شبکه‌های عصبی مصنوعی کاملاً قابل رقابت با مدل‌های رگرسیون هستند. بدین صورت که شاخص



- 1 Multi layer perceptron
- 2 Synaptic
- 3 Back propogation
- 4 Gradient decrease optimization
- 5 Learning rate
- 6 Momentum

7 این عبارت به مفهوم دوره است (یعنی یک بار داده ها به شبکه

داده شوند گوئیم یک epoch طی شده است)

- 8 Bias
- 9 Resilent Back Propogation
- 10 Coujogate
- 11 Hessian Matrix
- 12 Quasi-Newton
- 13 Levenberg-Marquardt
- 14 Jacobian
- 15 Dummy variables