

پیش‌بینی تأخیر قطارهای مسافری با استفاده از شبکه‌های عصبی

مسعود یقینی*، استادیار، دانشکده مهندسی راه آهن، دانشگاه علم و صنعت ایران، تهران، ایران
محمد مهدی خوشرفرتار، دانشجوی کارشناسی ارشد، دانشکده مهندسی راه آهن، دانشگاه علم و صنعت ایران، تهران، ایران
سیدمسعود سیدآبادی، دانشجوی کارشناسی ارشد، دانشکده مهندسی راه آهن، دانشگاه علم و صنعت ایران، تهران، ایران

E-mail: yaghini@iust.ac.ir

دریافت: ۸۹/۰۳/۲۹ - پذیرش: ۸۹/۰۷/۲۷

چکیده

هدف این مقاله ارائه مدلی مبتنی بر شبکه‌های عصبی پیشخور با دقت بالا برای پیش‌بینی تأخیر قطارهای مسافری در راه آهن جمهوری اسلامی ایران است. در این مقاله از سه روش مختلف برای ورودی شبکه‌های عصبی، شامل ورود به صورت اعداد حقیقی نرمال شده، تبدیل ورودیها به صورت دودویی و همچنین تبدیل ورودیها به رشته‌های دودویی استفاده شده است. برای یافتن ساختاری مناسب برای شبکه عصبی در مدل پیشنهادی از سه استراتژی مختلف استفاده شده است که به ترتیب استراتژی سریع، استراتژی پویا و استراتژی چندگانه نام دارند. داده‌های مورد استفاده در این پژوهش شامل تأخیر قطارهای مسافری راه آهن جمهوری اسلامی ایران از سال‌های ۱۳۸۳ تا ۱۳۸۷ است. برای آماده‌سازی داده‌ها، جهت حذف داده‌های متناقض و دارای نویز که همواره همراه با مجموعه داده‌های دنیای واقعی است، پیش پردازش کاملی صورت گرفته است، برای پیشگیری از بیش برآزش مدل با مشخصات داده‌های آموزشی، براساس روش اعتبارسنجی تقاطعی، داده‌ها به سه قسمت داده‌های آموزشی، داده‌های اعتبارسنجی و داده‌های آزمایشی تقسیم شدند. برای ارزیابی مدل پیشنهادی، نتایج حاصل از روشهای مختلف ورود داده و استراتژی‌های یافتن ساختار مناسب شبکه، با یکدیگر و همچنین با بعضی روشهای رایج پیش‌بینی از قبیل درختهای تصمیم و رگرسیون لجستیک چندجمله‌ای مقایسه شده‌اند. برای مقایسه مدل‌های مختلف شبکه‌های عصبی مورد استفاده در این مقاله زمان آموزش مدل، دقت پیش‌بینی بر روی داده‌های آزمایشی و اندازه شبکه در نظر گرفته شده است. برای انجام مقایسه‌ای دقیق میان تمامی مدل‌ها، نمودار زمان - دقت ترسیم شده است. نتایج به دست آمده نشان می‌دهند مدل پیشنهادی از دقت بیشتری برخوردار است.

واژه‌های کلیدی: شبکه عصبی، پیش‌بینی، تأخیر قطار مسافری، شرکت قطارهای مسافری رجا

۱. مقدمه

تأخیر را براساس تاریخ، علت و زمان تأخیر ثبت و تحلیل می‌کند. براساس اطلاعات کسب شده از این پایگاه داده، میانگین تأخیر سالیانه از سالهای ۱۳۸۳ تا ۱۳۸۷ به میزان ۱۸۱۷۴ ساعت در سال و میانگین تأخیر برای هر قطار ۳۰ دقیقه بوده است. با بررسی انجام شده بر روی ادبیات موضوع، تا کنون مقالات کمی در زمینه تحلیل و پیش‌بینی داده‌های تأخیر قطارهای مسافری ارائه شده است. کاری و کویسینسکی [Carey and Kwiecinski, 1994] یک روش آماری برای تخمین تأخیر

تأخیر قطارهای مسافری از مهم‌ترین چالشهای سیستم‌های ریلی در تمام دنیا به شمار می‌آید. بر طبق گزارش دفتر ممیزی ملی انگلستان تنها در ماه ژوئیه ۲۰۰۸ در انگلستان بیش از ۱۴ میلیون دقیقه تأخیر در قطارهای مسافری آن کشور رخ داده است که هزینه‌ای بالغ بر یک میلیارد یورو (به طور میانگین ۷۳ یورو در هر دقیقه) بر بخش مسافری راه آهن انگلستان تحمیل کرده است [National Audit Office, 2008]. با توجه به اهمیت این موضوع، راه آهن جمهوری اسلامی ایران، داده‌های مربوط به

پیش‌بینی تأخیر در پروازها مبتنی بر یادگیری ماشینی ارائه کرده‌اند. در این مطالعه نخست یک روش یادگیری غیرنظارتی برای داده‌های گردآوری شده از فرودگاه ارائه شده و سپس استانداردی برای هر کلاس از تأخیر به دست می‌آید. آنها با کلاسهای به دست آمده از تأخیرها، یک شیوه یادگیری نظارتی را بر روی این داده‌ها پیاده‌سازی می‌کنند تا مدلی برای شناسایی تأخیرهای هوایی ساخته شود.

در این مطالعه از شیوه خوشه‌بندی k-means برای یادگیری غیرنظارتی، و از روش درخت تصمیم، شبکه‌های عصبی پس انتشار و روش بیزی برای یادگیری نظارتی استفاده می‌شود. در مطالعه دیگری زونگلی و همکارانش [Zonglei, Jiandong and Tao, 2009] روشی برای پیش‌بینی تأخیر در پروازها ارائه می‌کنند که براساس پیش‌بینی انجام شده برای تأخیر و انتشار آن به فرودگاههای دیگر، یک سیستم هشدار دهنده ارائه می‌شود. این روش، محاسبات کمی دارد و پیش‌بینی را به سرعت انجام می‌دهد.

جیانلی و همکارانش [Jianli, Yuecheng and Jiandong, 2009] مدلی مبتنی بر روش اتوماتای یادگیری در حالت موازی برای پیش‌بینی تأخیر پروازها و انتشار این تأخیر ارائه کرده‌اند. لانگ و هاسان [Long and Hasan, 2009] روشی بهبود یافته برای پیش‌بینی تأخیر در پروازهای هوایی ارائه کرده‌اند. در مطالعه انجام شده یک مدل شبیه‌سازی ارائه گردیده که در آن شرایط مختلف نظیر هوای نامساعد در نظر گرفته شده است. ساختار مقاله به این صورت سازماندهی شده است: در بخش ۲ به تشریح داده‌های جمع‌آوری شده تأخیر قطارهای مسافری پرداخته می‌شود. در بخش ۳ نخست توضیح مختصری درباره نحوه بهنگام‌سازی وزنها در شبکه‌های عصبی داده می‌شود و سپس به توضیح در مورد روشهای مختلف برای یافتن ساختار مناسب برای مدل پیشنهادی پرداخته می‌شود. در بخش ۴ در ابتدا تقسیم‌بندی داده‌های موجود به مجموعه‌های آموزشی، آزمایشی و اعتبارسنجی^۴ توضیح داده می‌شود، سپس گسسته‌سازی تأخیرها، تبدیل ورودیها به رشته‌های دودویی و پس از آن نتایج به دست آمده از مدل پیش‌بینی، تشریح می‌گردد. در بخش ۵ به جمع‌بندی مطالب و نتیجه‌گیری نهایی و همچنین ارائه پیشنهادهایی برای تحقیقات آینده پرداخته می‌شود.

ضربه‌ای^۲ قطارها ارائه کرده‌اند. تأخیر ضربه‌ای به معنای تأخیری است که در یک قطار به وجود آمده و باعث ایجاد تأخیر در قطارهای بعد از خود می‌شود. هویسمن و بوچری [Huisman and Boucherie, 2001] یک مدل تحلیلی برای پیش‌بینی تأخیر قطارها با سرعتهای مختلف ارائه کرده‌اند. در این مدل برای اینکه بتوانند تأثیر حرکت قطارهای برنامه‌ریزی شده را نیز در نظر بگیرند، توجه خاصی به وابستگی میان زمانهای سیر و زمانهای حائل^۳ شده است. این مدل در راه آهن هلند برای کمینه کردن تأخیر قطارها مورد استفاده قرار گرفته است. پیترز و همکارانش [Peters (et al.), 2005] به منظور کمینه کردن تأخیر قطارها مدلی هوشمند ارائه کرده‌اند. این مدل به عنوان یک سیستم پیش‌بینی بلادرنگ برای بهینه‌سازی جدول زمانی در شبکه قطارها مورد استفاده قرار گرفته است.

یوان [Yuan, 2006] یک مدل احتمالی بهبود یافته برای تأخیر قطارها و انتشار تأخیر در ایستگاهها ارائه کرده است. مهم‌ترین نوآوری در این مطالعه، مدل احتمالی ارائه شده است که با دقت زیادی به پیش‌بینی تأخیر ضربه‌ای قطارها می‌پردازد.

یوان [Yuan, 2007] به بررسی وابستگی آماری در تأخیر قطارها و انتشار این تأخیرها پرداخته است. در این مطالعه سه نوع از وابستگیهای تأخیری فرمول شده و سپس یک مدل احتمالی انتشار تأخیرها بین دو ایستگاه متوالی ارائه شده است. بریگس و بک [Briggs and Beck, 2007] نشان می‌دهند که تأخیر در قطارهای بریتانیا از توزیع فوق آماری خاصی به نام توزیع نمایی q پیروی می‌کند. در این تحقیق، داده‌های زمان اعزام ۲۳ ایستگاه اصلی بریتانیا از سپتامبر سال ۲۰۰۵ تا اکتبر سال ۲۰۰۶ بررسی شده است.

دامن و همکارانش [Daamen, Goverde and Hansen, 2009] شیوه‌ای را برای مشخص کردن تأخیرهای ضربه‌ای به صورت خودکار و دقیق ارائه کرده‌اند. در این مطالعه حوادث و زمان انتظار برای اتصال قطار، به عنوان منابع ایجاد تأخیر در قطارها محسوب می‌شوند که توسط اعزام کننده‌ها ثبت شده‌اند. یکی از زمینه‌های مرتبط با موضوع این مقاله، پیش‌بینی تأخیر پرواز در فرودگاههاست. در ادامه به بررسی برخی از مطالعات انجام شده در این زمینه پرداخته می‌شود. زونگلی و همکارانش [Zonglei, Jiandong and Guansheng, 2008] روشی برای

۲. شناخت داده‌های تأخیر

- خرابی سالن، سانحه، خرابی خط، خرابی علائم و توقف غیر برنامه‌ای برای نماز.

در راه آهن ج.ا.ا داده‌های مربوط به تأخیرات قطارهای مسافری به صورت روزانه ثبت و ذخیره می‌شود، سپس در آخر هفته این داده‌ها با هم ادغام شده و داده‌های هفتگی تأخیرات را تشکیل می‌دهند. در پایان هر ماه نیز فایل‌های مربوط به داده‌های هفتگی با یکدیگر ادغام و فایل‌های ماهانه را ایجاد می‌کنند. در این تحقیق، فایل‌های ماهانه تأخیر از سال ۱۳۸۳ تا پایان سال ۱۳۸۷ گردآوری شده‌اند که تعداد رکوردها در هر سال بیانیگر تعداد قطارهای مسافری اعزامی در آن سال است. جدول ۱ نشان‌دهنده این اطلاعات به طور خلاصه است. همان‌طور که در جدول ۱ دیده می‌شود، سال ۱۳۸۶ بیشترین مجموع تأخیر را داشته است و بعد از آن سال ۱۳۸۵ و سپس سال ۱۳۸۷ بیشترین مجموع تأخیر را داشته‌اند. با توجه به جدول ۱، میانگین تأخیر برای هر قطار در سال ۱۳۸۶ برابر با / دقیقه بوده است که این میانگین از تقسیم مجموع تأخیر بر حسب دقیقه بر تعداد قطارهای اعزامی به دست آمده است. شکل ۱ نشان‌دهنده میانگین تأخیر برای هر قطار در سالهای مختلف است. از نظر میانگین، سال ۱۳۸۶ بیشترین مقدار میانگین تأخیر را داشته است. موضوع حائز اهمیت در این قسمت این است که با این که سال ۱۳۸۷ مجموع تأخیر نسبتاً زیادی داشته است، ولی در مقایسه با میانگین تأخیر مقدار خوبی داشته است.

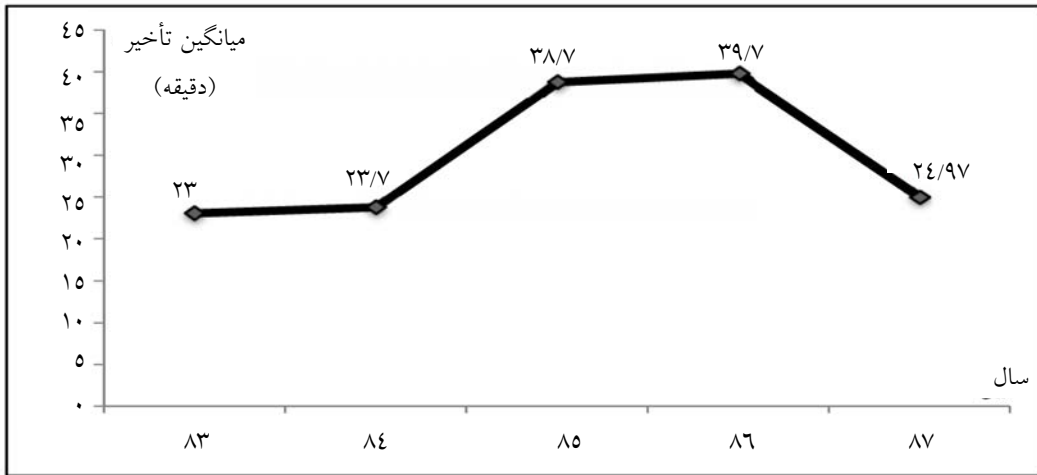
تأخیر قطارهای مسافری به این معناست که قطار طبق زمان‌بندی از پیش تعیین شده خود، حرکت نکرده باشد. تأخیر قطارها شامل توقفهای برنامه‌ای نیست، به این معنی که به عنوان مثال توقف قطار برای نماز یا توقف در ایستگاههای بین راهی برای سوار یا پیاده کردن مسافران جزء زمانهای تأخیر محسوب نمی‌شود. در واقع اگر قطار به هر علتی دیرتر از زمان مقرر به مقصد برسد، شامل تأخیر می‌شود. دلایل تأخیر در قطارهای مسافری می‌توانند به صورت زیر باشند:

- تأخیر در مبدا: اختلاف زمان حرکت واقعی قطار با زمان برنامه‌ریزی شده.
- تلاقی با قطارهای مسافری و باری: این مورد در زمانی که قطارهایی در مسیر مخالف در حرکت هستند اتفاق می‌افتد و به مدت زمانی می‌گویند که یک قطار منتظر می‌ماند تا یک بلاک خالی شود و آن قطار بتواند در آن بلاک سیر کند.
- نرسیدن قطار مسافری و باری: به مدت زمانی می‌گویند که یک قطار منتظر می‌ماند تا قطارهایی که در مسیر مشترک با این قطار هستند به این قطار رسیده و از این قطار بر طبق زمان‌بندی یا اولویت سبقت بگیرند.
- خرابی دیزل قطار: خراب شدن دیزل قطار مسافری.
- خرابی دیزل قطارهای دیگر: خرابی دیزل قطارهای دیگر که بر سیر این قطار تأثیر می‌گذارند.

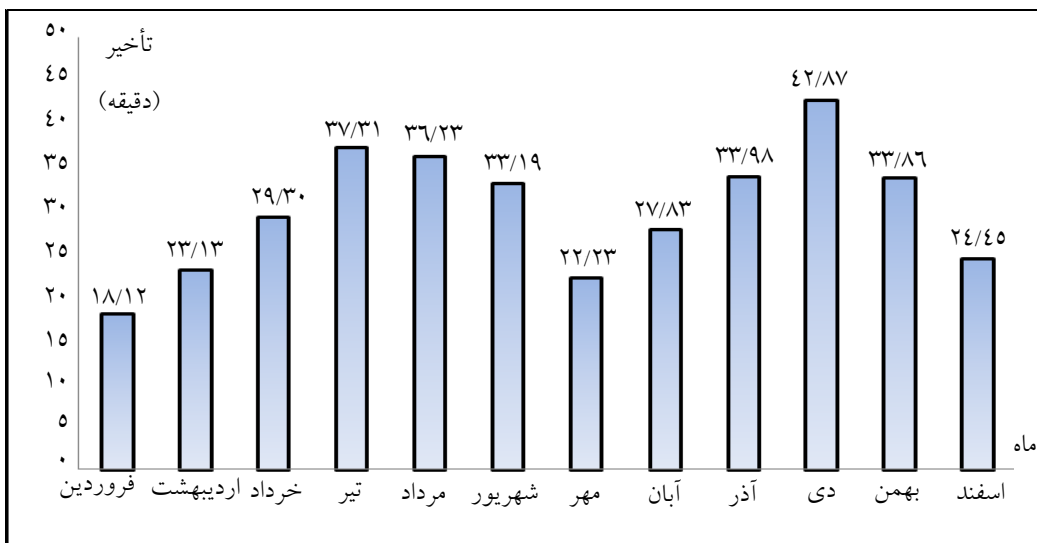
جدول ۱. خلاصه اطلاعات جمع‌آوری شده از تأخیر قطارهای مسافری [پایگاه داده قطارهای مسافری رجا، ۱۳۸۸]

| سال | تعداد قطار اعزامی | مجموع تأخیر (بر حسب دقیقه) | میانگین تأخیر برای هر قطار (بر حسب دقیقه) |
|-------|-------------------|----------------------------|-------------------------------------------|
| ۱۳۸۳ | ۲۸۹۲۸ | ۶۶۷۳۵۷ | ۲۳/۰ |
| ۱۳۸۴ | ۳۱۱۱۱ | ۷۳۷۱۸۲ | ۲۳/۷ |
| ۱۳۸۵ | ۳۴۱۶۶ | ۱۳۲۳۰۷۰ | ۳۸/۷ |
| ۱۳۸۶ | ۳۹۴۴۰ | ۱۵۶۷۴۵۶ | ۳۹/۷ |
| ۱۳۸۷ | ۴۶۳۳۷ | ۱۱۵۷۱۷۴ | ۲۴/۹ |
| مجموع | ۱۷۹۹۸۲ | ۵۴۵۲۲۳۹ | ۳۰/۰ (میانگین) |

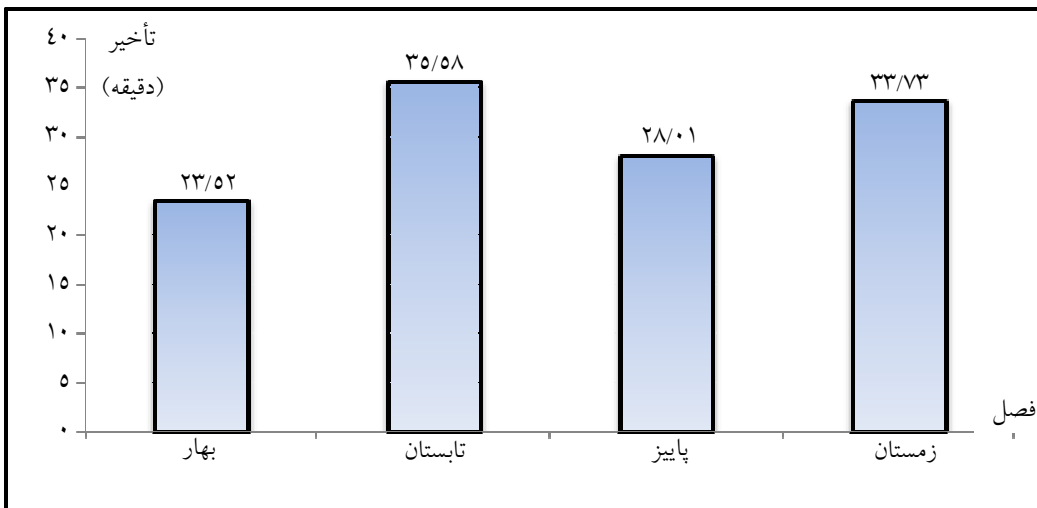
یقینی، خوشرفتار و سیدآبادی



شکل ۱. میانگین تأخیر برحسب دقیقه از سال ۱۳۸۳ تا ۱۳۸۷ (برحسب دقیقه)



شکل ۲. میانگین تأخیر ماهانه قطارهای مسافری از سال ۱۳۸۳ تا ۱۳۸۷ (برحسب دقیقه)



شکل ۳. میانگین تأخیر فصلی قطارهای مسافری از سال ۱۳۸۳ تا ۱۳۸۷ (برحسب دقیقه)

در نظر گرفته می‌شود، یعنی اگر $\eta(t-1) < \eta_{low}$ آنگاه $\eta(t) = \eta_{high}$. تابع کاهش لگاریتمی برای پارامتر نرخ یادگیری به صورت:

$$\eta(t) = \eta(t-1) \cdot \exp(\log(\eta_{low} / \eta_{high}) / d) \quad (3)$$

که در آن:

d مقداری است که توسط کاربر تنظیم می‌شود و به آن زوال η^{-1} گفته می‌شود. این فرآیند تا پایان زمان یادگیری ادامه می‌یابد.

مقدار خطا δ_{pj} برای لایه خروجی به صورت:

$$\delta_{pj} = (t_{pj} - o_{pj}) o_{pj} (1 - o_{pj}) \quad (4)$$

محاسبه می‌شود و برای دیگر لایه‌ها به صورت:

$$\delta_{pj} = o_{pj} (1 - o_{pj}) \sum_k \delta_{pk} w_{kj} \quad (5)$$

که t_{pj} مقدار خروجی مورد انتظار برای پیش‌بینی است. وزنهای شبکه بعد از پیش‌بینی خروجی هر نمونه بهنگام‌سازی می‌شوند.

۲-۳ ساختار شبکه عصبی برای مدل پیشنهادی

یکی از مباحث چالش برانگیز در بحث شبکه‌های عصبی، یافتن ساختار بهینه شبکه براساس داده‌های موجود است. برای غلبه بر این مشکل در این تحقیق از چندین استراتژی متفاوت استفاده شده است که در نهایت استراتژی با بیشترین دقت، به عنوان استراتژی برتر انتخاب شده است.

استراتژی سریع

در استراتژی سریع فقط یک شبکه عصبی از نوع پرسپترون سه لایه، به صورت مجزا آموزش داده می‌شود. این شبکه به طور پیش فرض دارای یک لایه مخفی است که در این لایه مخفی به تعداد $\max((\Theta_i + \Theta_o) / 20, 3)$ نرون مخفی وجود دارد، که Θ_i تعداد نرونها در لایه ورودی و Θ_o تعداد نرونها در لایه خروجی هستند. وزنهای شبکه با استفاده از الگوریتم آموزشی پس انتشار بهنگام‌سازی می‌شوند.

استراتژی پویا

در استراتژی پویا ساختار شبکه در طول فرآیند یادگیری تغییر می‌کند و نرونها به شبکه اضافه می‌شوند تا کارایی شبکه افزایش یافته و به دقت مطلوب برسد.

شکل ۲ نشان دهنده میانگین تأخیر ماهانه در قطارهای مسافری برای سالهای ۱۳۸۳ تا ۱۳۸۷ است که بیشترین میانگین تأخیر ماهانه با مقدار ۴۲/۸۷ دقیقه برای هر قطار مربوط به دی ماه و کمترین میانگین تأخیر در ماه با مقدار ۱۸/۱۲ دقیقه برای هر قطار، مربوط به فروردین ماه است. شکل ۳ میانگین فصلی تأخیرها را نشان می‌دهد، فصل تابستان با میانگین تأخیر ۳۵/۵۸ پرتأخیرترین فصل در این سالها بوده و همچنین فصل بهار با میانگین تأخیر ۲۳/۵۲، کمترین تأخیر را داشته است.

۳. مبانی مدل پیشنهادی

۱-۳ بهنگام‌سازی وزنها در شبکه عصبی پیشنهادی

شبکه عصبی مصنوعی یکی از تکنیکهای مهم در داده کاوی به شمار می‌آید که هم در یادگیری نظارتی و هم در یادگیری غیرنظارتی مورد استفاده قرار می‌گیرد. شبکه عصبی مورد استفاده در این مقاله، شبکه عصبی پیشخور^۵ است که به آن پرسپترون چند لایه^۶ نیز می‌گویند. برای آموزش شبکه از الگوریتم پس انتشار^۷ خطا به همراه اندازه حرکت^۸، استفاده شده است. تابع فعال‌سازی^۹ مورد استفاده برای این شبکه، تابع سیگموئید^{۱۰} به صورت زیر است:

$$\sigma(x) = 1 / (1 + e^{-x}) \quad (1)$$

در ابتدای یادگیری تمام وزنها در شبکه به صورت تصادفی در فاصله $-0.5 \leq w_{ij} \leq 0.5$ انتخاب می‌شوند. نحوه بهنگام‌سازی وزنها بعد از ورود هر نمونه از مجموعه داده‌ها به صورت زیر است:

$$\Delta w_{ij}(t+1) = \eta \delta_{pj} o_{pi} + \alpha \Delta w_{ij}(t) \quad (2)$$

که در آن:

پارامتر η نرخ یادگیری^{۱۱}، δ_{pj} خطای انتشار یافته، o_{pi} خروجی نرون i برای نمونه p پارامتر α اندازه حرکت و $\Delta w_{ij}(t)$ مقدار تغییر w_{ij} در تکرار قبلی است. مقدار α در حین فرآیند یادگیری ثابت است، ولی اندازه η در طول تکرارهای فرآیند یادگیری تغییر می‌کند. در ابتدا η به صورت یک مقدار اولیه داده می‌شود و سپس به صورت لگاریتمی به مقدار η_{low} کاهش می‌یابد، و هنگامی که مقدار η کمتر از مقدار η_{low} شد، آنگاه مقدار η_{high} را برای آن

این مرحله خود دارای دو زیر مرحله: یافتن ساختار مناسب و سپس آموزش شبکه نهایی است. برای یافتن ساختار مناسب، ابتدا شبکه‌ای با دو لایه مخفی ساخته می‌شود که هر یک دارای دو نرون هستند. نرخ یادگیری اولیه 0.05 و 0.9 α در نظر گرفته شده و شبکه به اندازه یک تکرار آموزش داده می‌شود و یک کپی از شبکه ایجاد کرده و یکی را راست و دیگری را چپ می‌نامیم، سپس به دومین لایه مخفی شبکه سمت راست یک نرون اضافه کرده و مجدداً هر دو شبکه به اندازه یک تکرار آموزش داده می‌شوند و مجموع خطا در هر دو شبکه به دست می‌آید. در صورتی که شبکه سمت چپ خطای کمتری داشته باشد آن را نگهداشته و یک نرون به اولین لایه مخفی شبکه سمت راست اضافه می‌شود. در صورتی که شبکه سمت راست خطای کمتری داشته باشد، یک نسخه از شبکه سمت راست به جای شبکه سمت چپ قرار داده می‌شود و مجدداً یک نرون به دومین لایه مخفی شبکه سمت راست اضافه می‌گردد. هر دو شبکه را به اندازه یک تکرار آموزش می‌دهیم و این فرآیند را تا رسیدن به شرط خاتمه تکرار می‌کنیم. برای تنظیم نرخ یادگیری در هر تکرار دو بردار محاسبه می‌شود، اولی بردار حرکت $M(t)$ بر مبنای تغییرات وزنها در یک تکرار و دومی بردار تغییر $C(t)$ که بر مبنای اندازه حرکت تکرار فعلی است. بردارهای $M(t)$ و $C(t)$ از روابط زیر محاسبه می‌شوند:

ساختار مناسب در استراتژی پویا برای شبکه عصبی، نوبت به آموزش شبکه نهایی براساس روش پس انتشار خطا می‌رسد. برای شروع آموزش نرخ یادگیری اولیه را برابر 0.02 و 0.9 α گرفته و بر این اساس شبکه را آموزش می‌دهیم.

استراتژی چندگانه

در استراتژی چندگانه تا رسیدن به شرط خاتمه، چندین شبکه به صورت شبه موازی با هم آموزش می‌بینند و شبکه‌ای که بیشترین دقت را دارد به عنوان شبکه نهایی انتخاب می‌شود. به این صورت که در ابتدا چندین شبکه با یک لایه مخفی ساخته می‌شوند که تعداد نرونها در لایه مخفی از ۳ تا حداکثر تعداد نرونهای لایه ورودی تغییر می‌کند. سپس به ازای هر شبکه با یک لایه مخفی چندین شبکه با دو لایه مخفی ایجاد می‌کنیم که تعداد نرونهای لایه مخفی اول درست به اندازه شبکه یک لایه‌ای است، ولی تعداد نرونهای لایه مخفی دوم متغیر و به ترتیب ۲، ۵، ۱۰ و ۱۷ است و به همین ترتیب حداکثر تا تعداد نرونهای لایه مخفی اول در لایه مخفی دوم می‌توانیم نرون داشته باشیم و با ساخت این شبکه‌ها، آنها را آموزش می‌دهیم [SPSS Clementine, 2008].

۴. پیش‌بینی تأخیر قطارهای مسافری

۴-۱ داده‌های آموزشی، آزمایشی و اعتبارسنجی

یکی از مشکلات رایج هنگام آموزش مدل‌های پیش‌بینی، تطبیق بیش از حد مدل با خصوصیات داده‌های آموزشی فعلی یا به عبارتی بیش‌برازش^{۱۳} است که این مسأله به پیش‌بینی نادرست و با دقت کمتر از حد انتظار برای مقادیر جدید منجر می‌شود. برای جلوگیری از چنین حالتی با توجه به حجم زیاد داده‌ها (به طور متوسط بیش از ۳۰,۰۰۰ نمونه برای هر سال)، با استفاده از روش اعتبارسنجی تقاطعی^{۱۴}، داده‌ها به سه بخش داده‌های آموزشی^{۱۵}، داده‌های آزمایشی^{۱۶} و داده‌های اعتبارسنجی تقسیم می‌شوند، سپس آموزش به صورت دوره‌ای متوقف شده و بعد از هر دوره آموزشی، شبکه با استفاده از زیرمجموعه اعتبارسنجی، مورد ارزیابی قرار می‌گیرد. با استفاده از این شیوه تشخیص شروع پیش‌برازش امکان‌پذیر خواهد شد.

$$M(t) = 2[W(t) - W(t-1)] \quad (6)$$

$$C(t) = 0.8.C(t-1) - M(t) \quad (7)$$

که در آن $W(t)$ بردار وزن تکرار فعلی و $W(t-1)$ بردار وزن تکرار قبلی است. بردار نسبت بزرگی این دو بردار را به صورت زیر تعریف می‌کنیم:

$$m(t) = \|M(t)\| / \|C(t)\| \quad (8)$$

این مقدار شاخص شتاب یادگیری است، در صورتی که این شاخص کمتر از $10 / \|C(t)\| + 1$ باشد یادگیری در حال کند شدن است و نرخ یادگیری در عدد $1/2$ ضرب می‌شود و اگر شاخص بیش از ۵ باشد یادگیری در حال شتاب گرفتن است و نرخ یادگیری در عدد $4/m(t)$ ضرب می‌شود. بعد از یافتن

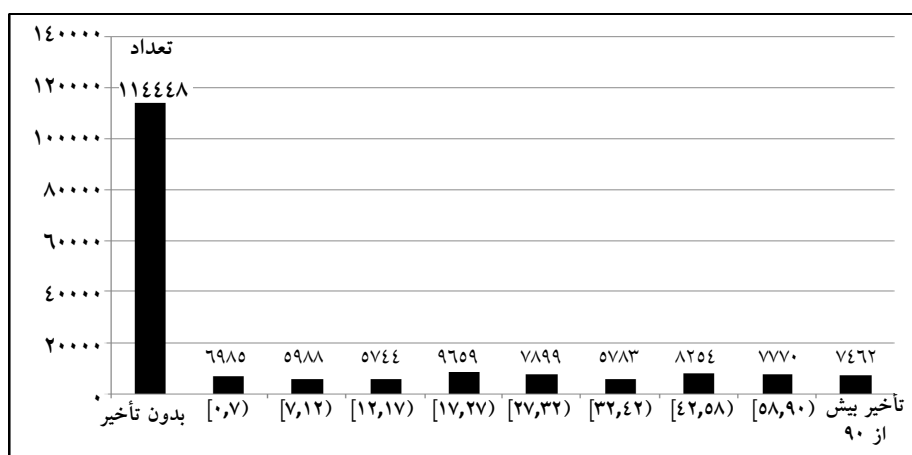
صحیح نسبت داده شده است. همچنین قطارها شامل ۱۰ محور حرکتی هستند که به هر یک از محورهای نیز یک عدد صحیح نسبت داده شده است. خروجی شبکه یکی از ده گروه مربوط به تأخیر قطارهاست که در بالا تشریح شد، به این صورت که در لایه خروجی ۱۰ گره در نظر گرفته می‌شود که خروجی هر کدام از گره‌ها یک مقدار باینری است و نمایش دهنده یکی از گروه‌های مربوط به تأخیر قطار است.

با استفاده از آزمون مربع کای پیرسن می‌توان میزان وابستگی میان متغیرهای ورودی از نوع طبقه‌ای را محاسبه کرد. در صورتی که مقدار آزمون مربع کای در فاصله $[0, \chi^2_{\alpha(r-1) \times (c-1)}]$ قرار گیرد شرط استقلال دو متغیر پذیرفته می‌شود، α سطح معنی‌دار بودن آزمون است که آن را ۰/۰۵ در نظر گرفته‌ایم، r تعداد سطرها و c تعداد ستونها است. در صورتی که احتمال عدم وابستگی کمتر از ۰/۰۵ باشد، وابستگی بین متغیرهای ورودی زیاد خواهد بود. در جدول ۲ مقادیر آزمون مربع کای پیرسن برای متغیرهای تأخیر، محور، روز، ماه، سال و مبدأ- مقصد آورده شده است.

در این مقاله از ۶۰٪ داده‌ها به عنوان داده‌های آموزشی، ۱۰٪ برای اعتبارسنجی و ۳۰٪ باقیمانده داده‌ها نیز به عنوان داده‌های آزمایشی بهره‌گیری شد.

۴-۲ نرونهای خروجی شبکه عصبی

در پایگاه داده‌ای موجود، تأخیر قطارهای مسافری به صورت اعداد حقیقی بر حسب دقیقه، ثبت شده است. این داده‌ها با استفاده از روش گسسته‌سازی به ده گروه مختلف تقسیم شده و هر یک از این گروه‌ها به عنوان یک نرون خروجی در شبکه عصبی تعریف می‌شوند. شکل ۴ فراوانی هر گروه را نشان می‌دهد. در نهایت مدل پیش‌بینی می‌تواند یک گروه را پیش‌بینی کند، یعنی مدل، حدود زمان تأخیر یک قطار را پیش‌بینی کند. متغیرهای مورد استفاده در پیش‌بینی شامل مبدأ- مقصد قطار، روز، ماه و سال حرکت (یعنی تاریخ حرکت) و محورهایی است که قطار بر روی آنها حرکت می‌کند. در پایگاه داده‌ای موجود ۲۷۸ مبدأ- مقصد وجود دارد که به هر مبدأ- مقصد یک عدد



شکل ۴. داده‌های تأخیر قطارها به صورت گسسته

جدول ۲. مقادیر آزمون مربع کای پیرسن برای متغیرهای تأخیر

| متغیرها | مقدار آزمون مربع کای پیرسن | درجه آزادی | احتمال عدم وابستگی |
|-----------------------|----------------------------|------------|--------------------|
| تأخیر، محور | ۱۹۹۹۷/۵۲۵ | ۸۱ | ۰ |
| تأخیر، ماه | ۵۹۳۴/۶۲۶ | ۹۹ | ۰ |
| تأخیر، سال | ۵۲۷۱/۱۴۵ | ۳۶ | ۰ |
| تأخیر، روز | ۴۱۱/۱۶۹ | ۲۷۰ | ۰ |
| تأخیر، مبدأ-مقصد قطار | ۶۱۶۵۵۲/۶ | ۲۴۹۳ | ۰ |

۳-۴ نرونهاي ورودی شبکه عصبی

در پیش‌بینی با استفاده از شبکه عصبی برای یافتن بهترین نوع ورودی، متغیرها به سه صورت به عنوان ورودی به شبکه داده شده است. حالت اول ۵ نرون به عنوان ورودی شبکه در نظر گرفته می‌شود که ورودی هر یک از این نرونها شامل مبدأ-مقصد، محور قطار، روز، ماه و سال است. برای حالت دوم در ابتدا مقادیر صحیح نسبت داده شده به مبدأ-مقصد، محور، روز، ماه و سال به صورت دودویی تبدیل شده سپس به تعداد $\log_2(\max \text{ number} + 1)$ متغیر جدید تولید می‌شود که $\max \text{ number}$ بزرگ‌ترین عدد صحیح در میان نمونه‌های متغیر مورد نظر است. در صورتی که مقدار بالا یک عدد صحیح نشد به نزدیک‌ترین عدد صحیح بیشتر از خودش تغییر داده می‌شود و در آخر مقادیر دودویی محاسبه شده برای هر متغیر در این متغیرهای جدید جایگزین شوند به صورتی که در هر متغیر تنها یک رقم از عدد دودویی قرار گیرد. در جدول ۳ برای مقدار ۵۰ این روند نمایش داده شده است (فرض کنیم بزرگ‌ترین عدد متغیر مربوطه ۷۰ باشد).

برای هر یک از حالت‌های متغیر طبقه‌ای یک متغیر باینری در نظر گرفته می‌شود. به طور مثال برای متغیر محور با ۱۰ حالت مختلف، ۱۰ متغیر باینری (برای هر حالت یک متغیر) در نظر گرفته می‌شود که در هر نمونه فقط یکی از ۱۰ متغیر مقدار یک می‌گیرد و به بقیه متغیرها مقدار صفر اختصاص می‌یابد. شکل ۷ نشان دهنده طرح کلی ساختار شبکه عصبی مورد استفاده در این مقاله با ورودی رشته‌های دودویی است. در این شکل X_1 تا X_9 نرونهاي ورودی دودویی برای متغیر مبدأ-مقصد است. در واقع هر یک از این نرونها، یکی از عناصر رشته هشت بیتی مربوط به متغیر مبدأ-مقصد را وارد شبکه عصبی می‌کند. به همین ترتیب،

D_1 تا D_5 نرونهاي ورودی دودویی برای متغیر روز، C_1 تا C_4 نرونهاي ورودی دودویی برای متغیر محور است. O_1 تا O_{10} نرونهاي دودویی خروجی شبکه برای پیش‌بینی هر یک از گروه‌های تأخیر است.

در شیوه ورودی دودویی، برای حضور یا عدم حضور هر یک از حالت‌های یک متغیر گسسته، یک نرون ورودی داریم، به عنوان مثال متغیری مانند روز می‌تواند تا ۳۱ حالت مختلف داشته باشد و بنابراین فقط برای این متغیر به ۳۱ نرون ورودی نیاز داریم که این مسأله به صورت نمایی باعث افزایش اندازه شبکه می‌شود، یا به عنوان مثالی دیگر متغیر مبدأ - مقصد در این پایگاه داده دارای ۲۷۸ حالت مختلف است، بنابراین برای این متغیر در حالت معمولی شبکه عصبی نیاز به ۲۷۸ نرون ورودی داریم، در حالی که شیوه رشته‌های دودویی فقط نیاز به ۹ نرون برای ورودی دارد. در صورتی که ورودی به صورت اعداد نرمال شده باشد اندازه شبکه به میزان قابل توجهی کاهش پیدا می‌کند به طوری که نرونهاي لایه ورودی در این حالت به ۵ عدد می‌رسد. جدول ۴ نشان دهنده مقایسه اندازه شبکه پیشنهادی شبکه‌های عصبی دودویی متداول، برای اولین استراتژی یافتن ساختار است.

لازم به یادآوری است که مقایسه ارائه شده در جدول ۳ فقط برای اولین استراتژی یافتن ساختار شبکه صورت گرفته است و برای دو استراتژی دیگر به دلیل عدم برخورداری از ساختاری ثابت، امکان انجام این محاسبات وجود ندارد و البته به دلیل استفاده از دو لایه مخفی، نهایتاً شبکه‌ها ساختاری بزرگ‌تر خواهند داشت. جدول ۵ نشان دهنده تعدادی نمونه از داده‌های مربوط به تأخیرات است، جدول ۶ نشانگر نمونه‌های تبدیل شده به اعداد صحیح است و جدول ۷ نمونه‌های به صورت رشته‌های دودویی را نشان می‌دهد.

جدول ۳. تبدیل عدد ۵۰ به صورت رشته دودویی

| $\log_2(70+1) = 6/14 \approx 7$ | $Bin(50) = (110010)_2$ | | | | | | |
|---------------------------------|------------------------|---------|---------|---------|---------|---------|---------|
| | متغیر ۱ | متغیر ۲ | متغیر ۳ | متغیر ۴ | متغیر ۵ | متغیر ۶ | متغیر ۷ |
| ۰۱۱۰۰۱۰ | ۰ | ۱ | ۱ | ۰ | ۰ | ۱ | ۰ |

پیش‌بینی تأخیر قطارهای مسافری با استفاده از شبکه‌های عصبی

جدول ۴. مقایسه اندازه ساختار روش پیشنهادی و روش متداول شبکه عصبی برای شیوه اول یافتن ساختار

| تعداد | ورودی عددی نرمال شده | ورودی دودویی | ورودی رشته‌دودویی |
|------------------------------------|----------------------|--------------|-------------------|
| نرون در لایه ورودی | ۵ | ۳۳۶ | ۲۵ |
| نرون در لایه مخفی اول | ۳ | ۱۸ | ۴ |
| نرون خروجی | ۱۰ | ۱۰ | ۱۰ |
| اتصال میان لایه ورودی و مخفی اول | ۱۵ | ۶۰۴۸ | ۱۰۰ |
| اتصالات میان لایه مخفی اول و خروجی | ۳۰ | ۱۸۰ | ۴۰ |
| کل اتصال‌های شبکه | ۴۵ | ۶۲۲۸ | ۱۴۰ |

جدول ۵. تعدادی از نمونه‌های پایگاه داده‌ای موجود

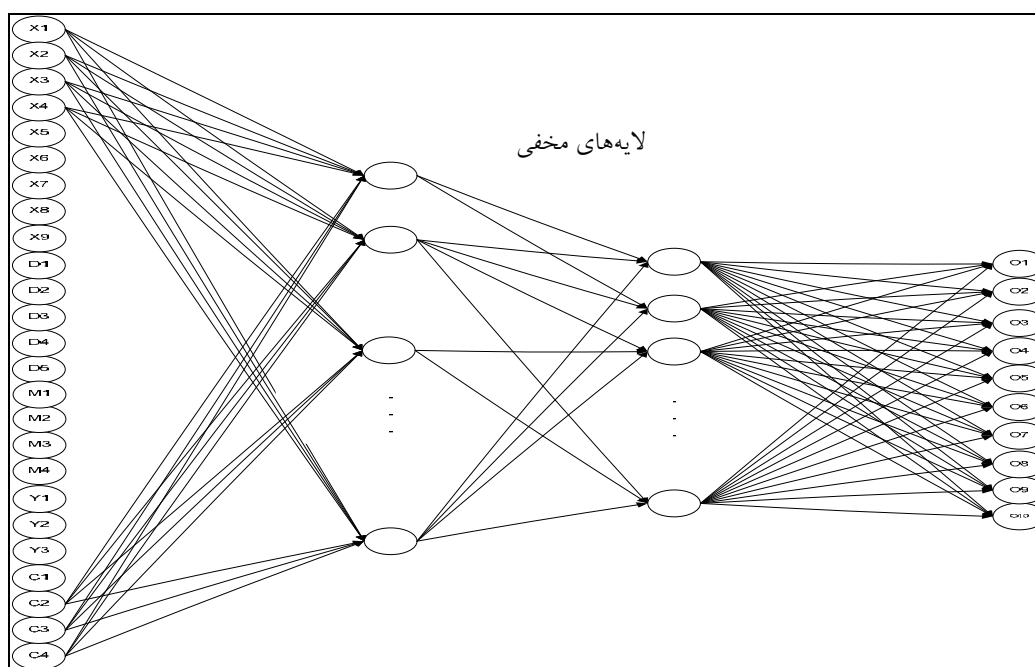
| مبدا- مقصد | تاریخ حرکت | محور | مجموع تأخیر |
|------------------|------------|-------------|-------------|
| تهران - زنجان | ۱۳۸۷/۰۵/۱۳ | آذربایجان | ۲۰۵ |
| بندرعباس - تهران | ۱۳۸۳/۱۱/۰۹ | جنوب شرق | ۸۵ |
| گرگان - تهران | ۱۳۸۴/۰۳/۱۰ | شمال | ۱۰ |
| اهواز - مشهد | ۱۳۸۵/۱۲/۲۳ | جنوب خراسان | ۰ |

جدول ۶. تبدیل نمونه‌ها به صورت عدد صحیح

| مبدا مقصد | روز | ماه | سال | محور | مجموع تأخیر (گروه‌بندی) |
|-----------|-----|-----|-----|------|-------------------------|
| ۷۰ | ۱۳ | ۵ | ۵ | ۲ | ۱۰ |
| ۲۰ | ۹ | ۱۱ | ۱ | ۸ | ۹ |
| ۲۴۱ | ۱۰ | ۰۳ | ۲ | ۱۰ | ۳ |
| ۱۵ | ۲۳ | ۱۲ | ۳ | ۷ | ۱ |

جدول ۷. تبدیل نمونه‌ها به رشته‌های دودویی

| مبدا- مقصد | روز | ماه | سال | محور | مجموع |
|------------|-------|------|-----|------|-----------|
| ۰۰۰۱۰۰۱۱۰ | ۰۱۱۰۱ | ۰۱۰۱ | ۱۰۱ | ۰۰۱۰ | ۰۰۰۰۰۰۰۰۱ |
| ۰۰۰۰۱۰۱۰۰ | ۰۱۰۰۱ | ۱۰۱۱ | ۰۰۱ | ۱۰۰۰ | ۰۰۰۰۰۰۰۱۰ |
| ۱۱۱۱۰۰۰۱ | ۰۱۰۱۰ | ۰۰۱۱ | ۰۱۰ | ۱۰۱۰ | ۰۰۱۰۰۰۰۰۰ |
| ۰۰۰۰۰۱۱۱۱ | ۱۰۱۱۱ | ۱۱۰۰ | ۰۱۱ | ۰۱۱۱ | ۱۰۰۰۰۰۰۰۰ |



شکل ۵. ساختار شبکه‌های عصبی با ورودی رشته‌های دودویی

که نتایج به دست آمده نشان می‌دهد، نمی‌توان یک روش خاص را به طور مطلق به عنوان بهترین روش انتخاب کرد، بلکه باید تحلیلی روی دقت، زمان و میزان حافظه مورد نیاز هر از یک روشها انجام داد. برای ارزیابی نتایج به دست آمده، از دو روش درخت تصمیم و رگرسیون لجستیک چند جمله‌ای^{۱۷} استفاده شده است. در پیش‌بینی با استفاده از درخت تصمیم و رگرسیون لجستیک چند جمله‌ای، مقادیر ورودی مدل پیش‌بینی به صورت عددی به مدل داده می‌شود. برای مقادیر خروجی به جای ده خروجی که در مدل پیشنهادی در نظر گرفته شد، در اینجا از یک خروجی (با مقادیر یک تا ده) استفاده می‌شود. الگوریتم C5.0 برای ساختن درخت تصمیم بکار گرفته شده است. در جدول ۹ نتایج حاصل از پیش‌بینی با درخت تصمیم و رگرسیون لجستیک آورده شده است. برای مقایسه آسان‌تر نتایج به دست آمده از روشهای مختلف، در جدول ۱۰ نتایج حاصل، از نظر دقت پیش‌بینی بر روی مجموعه داده‌های آزمایشی و همچنین زمان لازم برای آموزش مدل نشان داده شده است.

به هریک از روشها دو عدد صحیح نسبت داده شده است که یکی نشان دهنده رتبه مدل در میان بقیه از نظر دقت پیش‌بینی، و دیگری نشان دهنده رتبه مدل از نظر زمان آموزش است. در صورتی که زمان اجرا برای دو روش یکسان بوده است، عدد صحیح یکسانی نیز برای رتبه در نظر گرفته می‌شود.

۴-۴ مقایسه نتایج پیش‌بینی تأخیر با روش پیشنهادی و سایر روشهای پیش‌بینی

تمامی نتایج برای مدل پیشنهادی و همچنین سایر روشهای پیش‌بینی با استفاده از یک دستگاه رایانه با مشخصات Quad CPU 2.83 GHz 64 Bits با حافظه اصلی و همچنین با استفاده از نرم‌افزار [SPSS, 2008] Spss Clementine 12.0 به دست آمده است. در جدول ۷ نتایج حاصل از پیاده سازی مدل شبکه عصبی بر روی سه نوع مختلف ورود داده‌ها (یعنی روش عددی، روش دودویی و روش رشته دودویی) برای هریک از سه استراتژی مختلف یافتن ساختار (یعنی استراتژی سریع، استراتژی پویا و استراتژی چند گانه) ارائه شده است. از مجموع ۱۷۹۹۸۲ نمونه، تعداد ۱۰۷۹۱۸ نمونه برای آموزش، ۵۳۸۳۷ نمونه برای آزمایش و ۱۸۲۲۷ نمونه نیز برای اعتبارسنجی مدل مورد استفاده قرار گرفته است. در جدول ۸ مشاهده می‌شود که بالاترین دقت پیش‌بینی در داده‌های آزمایشی مربوط به داده‌های دودویی با استفاده از استراتژی سریع برای تعیین ساختار شبکه با مقدار ۹۲/۱۸٪ است. استفاده از داده‌های دودویی باعث افزایش اندازه شبکه شده و این موضوع باعث نیاز به حافظه بیشتر و همچنین زمان طولانی‌تر برای اجرای الگوریتم یادگیری شبکه است. کمترین دقت پیش‌بینی نیز مربوط به داده‌های عددی نرمال شده و با استراتژی چندگانه تعیین ساختار شبکه عصبی است. همان‌طور

پیش‌بینی تأخیر قطارهای مسافربری با استفاده از شبکه‌های عصبی

جدول ۸. نتایج مدل الگوریتم شبکه عصبی بر روی داده‌ها

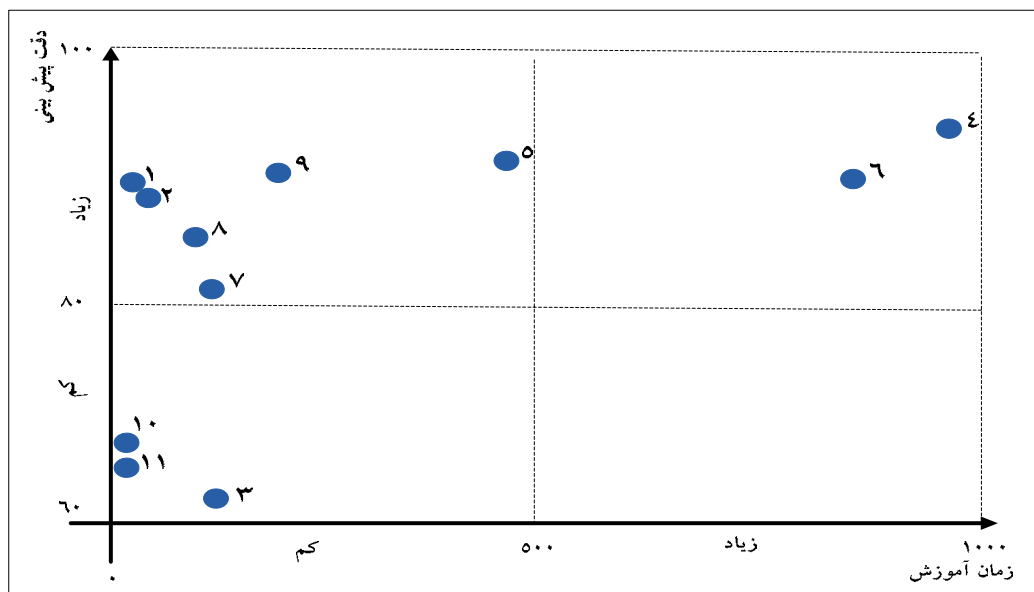
| داده‌های اعتبار سنجی | | داده‌های آزمایشی | | داده‌های آموزشی | | زمان (ثانیه) | استراتژی تعیین ساختار شبکه | نوع ورودی |
|----------------------|-------|------------------|-------|-----------------|-------|--------------|----------------------------|-------------|
| دقت | تعداد | دقت | تعداد | دقت | تعداد | | | |
| پیش‌بینی | درست | پیش‌بینی | درست | پیش‌بینی | درست | | | |
| ۹۰/۸۱ | ۱۶۵۵۳ | ۹۰/۸۰ | ۴۸۸۸۲ | ۹۰/۸۱ | ۹۸۰۰۰ | ۱۵ | سریع | عددی |
| ۹۰/۷۴ | ۱۶۵۳۸ | ۹۰/۷۰ | ۴۸۸۲۸ | ۹۰/۶۸ | ۹۷۸۵۹ | ۲۰ | پویا | |
| ۶۲/۳۵ | ۱۱۳۶۵ | ۶۲/۴۰ | ۳۳۵۹۴ | ۶۲/۳۰ | ۶۷۲۳۵ | ۵۷ | چندگانه | |
| ۹۲/۱۶ | ۱۶۷۹۹ | ۹۲/۱۸ | ۴۹۶۲۹ | ۹۲/۱۵ | ۹۹۴۴۹ | ۹۶۳ | سریع | دودویی |
| ۹۱/۴۹ | ۱۶۶۷۵ | ۹۱/۴۹ | ۴۹۲۵۶ | ۹۱/۴۴ | ۹۸۶۷۶ | ۴۵۷ | پویا | |
| ۹۱/۱۰ | ۱۶۶۰۶ | ۹۱/۱۶ | ۴۹۰۷۷ | ۹۱/۰۵ | ۹۸۲۵۷ | ۸۲۱ | چندگانه | |
| ۸۸/۸۲ | ۱۶۱۹۱ | ۸۸/۸۷ | ۴۷۸۴۳ | ۸۹/۰۹ | ۹۶۱۱۰ | ۴۱ | سریع | رشته دودویی |
| ۸۹/۹۸ | ۱۶۳۹۳ | ۸۹/۹۸ | ۴۸۴۴۴ | ۹۰/۴۸ | ۹۷۶۵۲ | ۲۸ | پویا | |
| ۹۰/۴۸ | ۱۴۹۲ | ۹۱/۱۲ | ۴۹۵۰۴ | ۹۱/۲۷ | ۹۸۴۹۳ | ۶۳ | چندگانه | |

جدول ۹. نتایج پیش‌بینی با درخت تصمیم و رگرسیون لجستیک

| داده‌های اعتبار سنجی | | داده‌های آزمایشی | | داده‌های آموزشی | | زمان (ثانیه) | الگوریتم |
|----------------------|------------|------------------|------------|-----------------|------------|--------------|----------------|
| دقت پیش‌بینی | تعداد درست | دقت پیش‌بینی | تعداد درست | دقت پیش‌بینی | تعداد درست | | |
| ۶۳/۸۳ | ۱۱۶۳۵ | ۶۴/۳۱ | ۳۴۶۲۳ | ۶۵/۴۲ | ۷۰۶۰۰ | ۲ | درخت تصمیم |
| ۶۳/۱۷ | ۱۱۵۱۴ | ۶۳/۶۰ | ۳۴۲۴۱ | ۶۳/۷۱ | ۶۸۷۵۵ | ۲۵ | رگرسیون لجستیک |

جدول ۱۰. مقایسه نتایج پیش‌بینی از نظر سرعت و دقت بر روی داده‌های آزمایشی

| رتبه در دقت | رتبه در زمان اجرا | دقت پیش‌بینی | زمان آموزش (ثانیه) | کد روش | استراتژی تعیین ساختار | نوع ورودی | روش |
|-------------|-------------------|--------------|--------------------|--------|-----------------------|-------------|----------------|
| ۵ | ۲ | ۹۰/۸۰ | ۱۵ | ۱ | سریع | عددی | شبکه عصبی |
| ۶ | ۳ | ۹۰/۷۰ | ۲۰ | ۲ | پویا | | |
| ۱۱ | ۷ | ۶۲/۴۰ | ۵۷ | ۳ | چندگانه | | |
| ۱ | ۱۱ | ۹۲/۱۸ | ۹۶۳ | ۴ | سریع | دودویی | |
| ۲ | ۹ | ۹۱/۴۹ | ۴۵۷ | ۵ | پویا | | |
| ۴ | ۱۰ | ۹۱/۱۶ | ۸۲۱ | ۶ | چندگانه | | |
| ۸ | ۶ | ۸۹/۰۹ | ۴۱ | ۷ | سریع | رشته دودویی | |
| ۷ | ۵ | ۹۰/۴۸ | ۲۸ | ۸ | پویا | | |
| ۳ | ۸ | ۹۱/۲۷ | ۶۳ | ۹ | چندگانه | | |
| ۹ | ۱ | ۶۴/۳۱ | ۲ | ۱۰ | - | عددی | |
| ۱۰ | ۴ | ۶۳/۶۰ | ۲۵ | ۱۱ | - | عددی | رگرسیون لجستیک |



شکل ۸. نمودار زمان آموزش - دقت پیش‌بینی

شبکه عصبی با ورودی دودویی برای تمامی ساختارها جزء روشهای با دقت بسیار ولی زمان آموزش بسیار طولانی است.

۵. نتیجه‌گیری

در این مقاله، مدلی مبتنی بر شبکه‌های عصبی با دقت بسیار برای پیش‌بینی تأخیر قطارهای مسافری در راه آهن ج.ا.ا ارائه شده است. در مدل ارائه شده، از ۳ روش مختلف برای ورودی داده‌ها به شبکه عصبی شامل ورود به صورت اعداد حقیقی نرمال شده، تبدیل ورودیها به صورت دودویی و همچنین تبدیل ورودیها به صورت رشته‌های دودویی استفاده شده است. همچنین برای تعیین ساختار مناسب شبکه از ۳ استراتژی مختلف استفاده شده است. برای پیش‌بینی تأخیر قطارهای مسافری، اطلاعات سالهای ۱۳۸۳ تا ۱۳۸۷ به کار گرفته شده است. برای ارزیابی مدل پیشنهادی، نتایج به دست آمده با نتایج حاصل از دو روش دیگر شامل درخت تصمیم و رگرسیون لجستیک مقایسه شده است. مقایسه نتایج به دست آمده نشان می‌دهد مدل پیشنهادی از دقت بیشتری برخوردار است. با استفاده از مدل ارائه شده می‌توان پیش‌بینی دقیقی از میزان تأخیر قطارهای مسافری داشته و نسبت به برنامه‌ریزیهای مناسب اقدام کرد. بنابراین می‌توان امیدوار بود که قطارهای مسافری، سیری با حداقل تأخیر داشته باشند.

برای یافتن شیوه‌ای مناسب برای مقایسه عادلانه میان مدل‌های مختلف، نمودار زمان آموزش - میزان دقت در پیش‌بینی در شکل (۸) ارائه شده است. در این نمودار، محور افقی به عنوان زمان آموزش مدل و محور عمودی به عنوان میزان دقت پیش‌بینی روی مجموعه داده‌های آموزشی در نظر گرفته شده است. با استفاده از این نمودار می‌توان جایگاه هر یک از روشهای استفاده شده را براساس میزان دقت و زمان یادگیری نسبت به سایر روشها مشاهده کرد. بر اساس پراکندگی نقاط بر روی این نمودار می‌توان روشهای پیش‌بینی را به ۳ گروه تقسیم کرد.

با توجه به شکل ۶، روش شبکه عصبی با ورودیهای دودویی برای تمام ساختارها و روش شبکه عصبی با ورودی عددی برای ساختار اول و دوم در زمان کوتاه، دقت بسیاری در پیش‌بینی دارند. همچنین روش درخت تصمیم، رگرسیون لجستیک و روش شبکه عصبی با ورودی عددی برای ساختار سوم، زمان آموزش کوتاه و دقت کمتری دارند. هر سه روش ساختاری برای مدل شبکه عصبی با ورودی به صورت رشته‌های دودویی در ناحیه اول یعنی بهترین ناحیه قرار دارند و این مطلب نشان دهنده کارایی شیوه پیشنهادی بر روی داده‌های تأخیر قطارهای مسافری راه آهن ج.ا.ا است. شیوه شبکه عصبی با ورودی دودویی معمولی با وجود دقت پیش‌بینی بالا، به علت زمان طولانی یادگیری، در ناحیه دوم واقع شده است و سایر روشها به دلیل دقت پیش‌بینی کمتر در ناحیه چهارم واقع می‌شوند. روش

- Huisman, T. and Boucherie, R. (2001) "Running times on railway sections with heterogeneous train traffic", *Transportation Research*, Vol. 35, March, pp. 271-292.

- Jianli, D., Yuecheng, Y. and Jiandong, W. (2009) "A model for predicting flight delay and delay propagation based on parallel cellular automata", *International Colloquium on Computing, Communication, Control, and Management*, Vol. 1, 29, September, pp.70 – 73.

- Long, D. and Hasan, S. (2009) "Improved prediction of flight delays using the LMINET2 System-Wide simulation model", *Aviation Technology, Integration and Operations Conference (ATIO)*, 23, September.

- National Audit Office (2008) "Reducing passenger rail delays by better management of incidents", HC 308 Session 2007-2008, Report by the Comptroller and Auditor General.

- Peters, J., Emig, B., Jung, M. and Schmidt, S. (2005) "Prediction of delays in public transportation using neural networks", *International Conference on Computational Intelligence for Modeling, Control and Automation, and International Conference on Intelligent Agents*, Vol. 2, pp. 92-97.

- SPSS (2008), *Clementine 12.0 Algorithms Guide*.

- Yuan, J. (2006) "Stochastic modeling of train delays and delay propagation in stations", Ph.D. dissertation, Delft University of Technology, Faculty of Civil Engineering and Geosciences, Department of Transportation and Planning.

- Yuan, J. (2007) "Dealing with stochastic dependence in the modeling of train delays and delay propagation", *International Conference on Transportation Engineering*.

- Zonglei, L., Jiandong, W. and Guansheng, Z. (2008) "A new method to alarm large scale of flights delay based on machine learning", *International Symposium on Knowledge Acquisition and Modeling*, 30 December, pp. 589 – 592.

- Zonglei, L., Jiandong, W. and Tao, X. (2009) "A new method for flight delays forecast based on the recommendation system", *Computing Communication Control and Management*, Vol. 1, p.p. 46-49.

زمینه‌های مختلفی برای ادامه این مطالعه می‌توان معرفی کرد. برای بهبود دقت پیش‌بینی می‌توان از روش‌های فرا ابتکاری نظیر الگوریتم ژنتیک یا تبرید شبیه‌سازی شده برای یافتن ساختار بهینه شبکه عصبی استفاده کرد. همچنین برای افزایش سرعت یادگیری شبکه، به ویژه برای شبکه‌هایی که ورودی دودویی دارند می‌توان از الگوریتم‌های فرا ابتکاری استفاده کرد.

۶. پی‌نوشت‌ها

1. National Audit Office (NAO)
2. Knock on Delay
3. Buffer Time
4. Validation Data
5. Feed Forward Neural Networks
6. Multilayer Perceptrons (MLP)
7. Back Propagation Algorithm
8. Momentum
9. Activation or Transfer Function
10. Sigmoid Function
11. Learning Rate
12. Decay
13. Overfitting
14. Crossed Validation
15. Training Data
16. Test Data
17. Multinomial Logistic Regression

۷. مراجع

- شرکت قطارهای مسافری رجا (۱۳۸۸) "پایگاه داده‌های تأخیرات قطارهای مسافری"، تهران: شرکت قطارهای مسافری رجا.

- Briggs, K. and Beck, C. (2007) "Modeling train delays with q-exponential functions", *Statistical Mechanics and its Applications*, 15, May, Vol. 378, pp. 498-504.

- Carey, M. and Kwiecinski, A. (1994) "Stochastic approximation to the effects of headways on knock-on delays of trains", *Transportation Research*, Vol. 28, August, pp. 251-267.

- Daamen, W., Goverde, R. and Hansen, I. A. (2009) "Non-discriminatory automatic registration of knock-on train delays", *Networks and Spatial Economics*, Vol. 9, No. 23, November, pp.47-61.