پیش‌بینی کوتاه‌مدت تعداد مسافر خطوط اتوبوس شهر تهران: نمونه موردی خط پل کریمخان- میدان جمهوری

سید احسان سید ابریشمی (مسئول مکاتبات)، استادیار، دانشکده عمران و محیط زیست، دانشگاه تربیت مدرس، تهران

وجیهه امینی، دانش آموخته کارشناسی ارشد، دانشکده عمران و محیط زیست، دانشگاه تربیت مدرس، تهران

مریم ایرانمنش، دانش آموخته کارشناسی ارشد، دانشکده عمران و محیط زیست، دانشگاه تربیت مدرس، تهران

علی محدث دیلمی، دانش آموخته دوره دکتری دانشکده برق و مهندسی کامپیوتر، دانشگاه تربیت مدرس، تهران

E-mail: seyedabrishami@modares.ac.ir

دریافت: ۱۳۹۶/۰۵/۱۸ پذیرش: ۱۳۹۶/۱۰/۰۶

چکیده

در دهه‌های اخیر افزایش جمعیت شهری و توسعه شهرنشینی از یک سو و عدم تکافوی حمل‌ونقل عمومی در مقابل تقاضای روز افزون سفرهای درون شهری از سوی دیگر، منجر به افزایش استفاده از خودروی شخصی، در تهران شده است. بنابراین تغییر در سیاست‌های حمل‌ونقل شهری و تلاش در جهت توسعه حمل‌ونقل عمومی، به ویژه اتوبوس، یکی از مهم‌ترین اقدامات در حوزه حمل‌ونقل شهری بوده و نیازمند بررسی‌های لازم جهت توسعه و کارآمد نمودن این شیوه‌ی سفر است. بررسی‌ها نشان می‌دهد، برنامه‌ریزی برای استفاده بهینه از زیرساخت‌های اتوبوسرانی و بکارگیری راهکارهای افزایش کارایی این سیستم در دنیا، نیازمند وجود اطلاعات در زمینه زیرساخت و تقاضای سفر خطوط اتوبوس است. بر این اساس، انجام مطالعاتی جهت برآورد تعداد مسافر خطوط اتوبوس در شهر تهران، که در حال حاضر تعیین تعداد مسافر هر خط با روش سنتی آماربرداری انجام شده و بسیار هزینه‌بر بوده، همچنین برنامه‌ریزی عملیاتی خطوط اتوبوسرانی که به صورت تجربی توسط مدیر خط انجام شده و هیچگونه برنامه‌ریزی مبتنی بر شرایط موجود برای آینده وجود ندارد، لازم و ضروری است. بدین ترتیب در این مطالعه با استفاده از اطلاعات AFC (داده‌های ثبت‌شده حاصل از تراکنش‌های کارت‌بلیت‌های هوشمند برای پرداخت کرایه) موجود در سیستم اتوبوسرانی شهر تهران و ساخت مدل مناسب، تعداد مسافر خطوط اتوبوس برای آینده کوتاه‌مدت پیش‌بینی گردید. به این منظور ابتدا اطلاعات AFC مربوط به هر خط اتوبوس مرتب شده و بر اساس آن مدل ساریما و شبکه عصبی (پرسپترون چند لایه) پرداخت می‌شود. نتایج مدل نشان می‌دهد، مدل پروسپترون چند لایه از لحاظ شاخص‌های خطا در پیش‌بینی تعداد مسافر خط از مدل ساریما برتر بوده و روش مناسب‌تری جهت برآورد تعداد مسافر خطوط اتوبوس است.

واژه های کلیدی: تقاضای مسافر خطوط اتوبوس، پرسپترون چند لایه، پیش‌بینی کوتاه‌مدت، ساریما، شبکه عصبی.

۱. مقدمه

یکی از بخش‌های مهم حمل‌ونقل که عهده‌دار جابجایی تعداد زیادی از افراد و اقشار جامعه است و سفرهای کاری و دیگر سفرهای روزمره افراد توسط این خدمت شهری انجام می‌گیرد، سیستم اتوبوسرانی است. بررسی‌ها نشان می‌دهد، برنامه‌ریزی برای استفاده بهینه از زیرساخت‌های حمل‌ونقل همگانی از جمله اتوبوسرانی شامل چهار مرحله، طراحی شبکه[۱]، توسعه جدول زمانی[۲]، زمان‌بندی وسایل نقلیه[۳] و زمان‌بندی کارکنان[٤] است و هر مرحله، نیازمند وجود اطلاعات در زمینه زیرساخت، حجم و زمان-بندی تقاضای موجود در این بخش است [Ceder, 2007]. بنابراین پیش‌بینی دقیق تقاضای سفر با اتوبوس، اهمیت بسیار زیادی در برنامه‌ریزی کوتاه‌مدت اتوبوسرانی داشته و باعث بهبود عملکرد و اعتماد به خدمات این شیوه از حمل‌ونقل همگانی می‌شود.

اهمیت پرداختن به این موضوع در تهران، کلان‌شهری با بیش از ۸ میلیون سکنه، که طبق آمار جمعیت این شهر در طول روز از ۱۲ میلیون فراتر رفته و تعداد سفرهای روزانه به ۱۷/۷ میلیون سفر می‌رسد، دوچندان است. این تعداد سفر، منشأ بسیاری از مشکلات تهران شده که برای نمونه می‌توان به مصرف روزانه بیش از ۱۲ میلیون لیتر بنزین توسط وسایل نقلیه که منجر به افزایش آلودگی محیط زیست و کاهش ضریب ایمنی و سلامتی شهروندان شده، اشاره نمود. از طرفی بر اساس آخرین مطالعات مبدأ- مقصد صورت گرفته در شهر تهران در سال ۱۳۸۳، تقریبا ۳۷/۹۵ درصد از سفرهای انجام شده در این شهر با استفاده از خودروی شخصی، ۲۱/۸٤ درصد تاکسی و مسافربر، ۵/٤۸ درصد سرویس، ۹/۱٦ درصد موتورسیکلت، ۲/۷۷ درصد وانت، ۱/۱۳ درصد مینی‌بوس و ۲۱/٦۷ درصد به وسیله حمل‌ونقل همگانی (مترو و اتوبوس) انجام می‌گیرد [O-D Tehran Matrice, 2004]. بنابراین یکی از راهکارها برای حل مشکلات این شهر، افزایش سهم استفاده از حمل‌ونقل همگانی بخصوص اتوبوس است که این امر نیازمند برنامه‌ریزی و دسترسی به اطلاعات دقیق بوده و مهم‌ترین اطلاعات مورد نیاز در این زمینه برآورد دقیق تقاضای سفر با اتوبوس برای خطوط مختلف است.

بررسی‌ها نشان می‌دهد، تاکنون مطالعات اندکی در زمینه پیش‌بینی کوتاه‌مدت تقاضای سفر اتوبوس در جهان انجام شده، برای نمونه، ژو، سان و چن [Xue, Sun and Chen, 2015] جهت تخمین پیش‌بینی کوتاه‌مدت سفر با اتوبوس از داده‌های سری زمانی تراکنش‌های کارت‌های هوشمند در یکی از خطوط شلوغ شهر شنژن[۵] چین، طی مدت ٤ ماه (آگوست تا نووامبر ۲۰۱۳) استفاده نمودند. برای تحلیل داده‌های سری زمانی، نرم‌افزار ای‌ویوس[٦] بکار گرفته شد و به منظور ساخت مدل مناسب برای هر نوع داده، تست ایستا[۷] انجام گرفت. در این مطالعه سه مدل پیش‌بینی هفتگی، روزانه و پانزده دقیقه‌ای برای بررسی ویژگی‌های متفاوت سری‌های زمانی پرداخت شد و با توجه تناوب ثابت سری زمانی هفتگی مدل ارما[۸]، تغییرات دوره‌ای سری زمانی روزانه مدل ساریما[۹] و نوسان گسترده تقاضای سفر در بازه‌های زمانی پانزده دقیقه طی یک روز، مدل اریما- گرچ[۱۰] (به تفکیک روزهای کاری و روزهای پایانی هفته) انتخاب شد. با توجه به متفاوت بودن مشخصات سری زمانی هفتگی، روزانه و پانزده دقیقه‌ای، ساخت مدل واحد که پیش‌بینی دقیقی از ویژگی‌های موقعیت‌های مختلف داشته باشد، امکان‌پذیر نبوده، بنابراین در این مطالعه از الگوریتم مدل چندگانه تعاملی[۱۱] استفاده شد که اطلاعات متفاوت و مدل‌های مختلف را ترکیب نموده و خطاهای ناشی از مدل واحد را حذف می‌نماید و بر اساس نتایج، عملکرد بهتری نسبت به مدل‌های جداگانه دارد. همچنین، سیپریچ، کنکنی و کیلیانوا [Cyprich, Konecny and Kilianova, 2013] به معرفی روشی برای مدلسازی تقاضای سفر با استفاده از تئوری اطلاعات سری زمانی تک‌متغیره و همچنین ارزیابی عملکرد رویکردهای مختلف پیش‌بینی پرداختند. اطلاعات مورد نیاز برای انجام این مطالعه، از شمارش کرایه کامل برای هر مسافر که توسط متصدیان در شهر زیلینا[۱۲] اسلواکی[۱۳] جمع‌آوری شده، تهیه و نرم‌افزار مورد استفاده سس[۱٤] است. به منظور پیش‌بینی تقاضای مسافران سیستم اتوبوس حومه شهری، از مدل‌های باکس- جنکینس، یکنواخت نمایی، رگرسیون خطی چندگانه، سارما و مدل اریمای[۱۵] تغییر شکل یافته لگاریتمی استفاده و مشاهده شد، عملکرد مدل اریمای تغییر شکل یافته بهتر از سایر مدل‌ها است.

ما و همکاران [Maa et al. 2014] با استفاده از داده‌های تراکنش ثبت شده کارت‌های هوشمند یکی از خطوط شهر جینان[16] چین که در طول یک سال جمع‌آوری شده به پیش‌بینی تقاضای کوتاه‌مدت مسافران اتوبوس پرداختند. برای مدلسازی، اطلاعات سری زمانی به صورت سه الگوی هفتگی، روزانه و ساعتی در نظر گرفته شد و برای اطلاعات هفتگی با توجه به توزیع ثابت با تغییرات غیرعادی مدل ار[17]، برای اطلاعات روزانه با تغییرات دوره‌ای، مدل ساریما و برای اطلاعات ساعتی، مدل اریما انتخاب و برای روزهای کاری و پایانی هفته به صورت جداگانه پرداخت شد. با توجه به ویژگی‌های خاص الگوی تقاضا در بازه‌های زمانی مختلف، نیاز به ساخت مدل واحد برای پیش‌بینی تقاضا وجود دارد که از مدل شبکه عصبی مصنوعی با الگوی هیبریدی[18] و مدل چندگانه تعاملی با الگوی هیبریدی استفاده گردید و مشاهده شد، مدل چندگانه تعاملی نسبت به سایر مدل‌ها، پیش‌بینی بهتری از نظر دقت، قابلیت پیش‌بینی و پیچیدگی مدل ارائه می‌کند. ژوآ و همکاران [Zhoua et al. 2016]، به ارائه یک سیستم پیش‌بینی تقاضای سفر اتوبوس با استفاده از اطلاعات سیستم واقعی اتوبوس پرداختند. پایگاه داده مطالعه جاری با استفاده از اطلاعات ارسالی ۱۳۲۶ اتوبوس مجهز به دستگاه جی‌پی‌اس[19] در ٤١٦ ایستگاه اتوبوس تهیه شده و مدت زمان برداشت اطلاعات یک دوره ۲۲ هفته‌ای بدون وقفه، در شهر یانتای[20] چین است. برای ساخت مدل پیش‌بینی مسافر برای هر اتوبوس در ایستگاه و زمان مشخص، بر اساس موقعیت تاریخی جی‌پی‌اس و اطلاعات سرویس مسافر سوار و پیاده شده در هر ایستگاه برای دوره زمانی کوتاه، نمودار سری زمانی هر ایستگاه اتوبوس که شامل تعداد تقاضای مسافر است، ساخته شد. سپس سه مدل مجزا تغییرات زمانی پوآسون[21] برای حل مشکل ناهمگونی اطلاعات، مدل وزن‌دهی تغییرات زمانی پوآسون[22] برای حل مشکلات تغییرات فصلی (نظیر تعطیلات و تغییرات آب‌وهوایی) و مدل اریما به منظور امکان پیش‌بینی تغییرات دوره‌ای غیر عادی، ارائه شد. در نهایت، یک چارچوب کلی وزنی[23] برای پیش‌بینی تعداد مسافر با یک مدل واحد با استفاده از اطلاعات سری زمانی، ارائه گردید. نتایج مطالعه نشان داد، عملکرد مدل کلی وزنی نسبت به سه مدل پوآسون، پوآسون وزنی و اریما بهتر است.

همچنین پیندو [Pinedo, 2014] به تخمین تعداد مسافر اتوبوس هر ایستگاه و مشخص نمودن عوامل محیط شهری[24] اثرگذار بر تعداد مسافران هر ایستگاه پرداخت. مطالعه موردی، شهر مادرید اسپانیا بوده و یکی از مهم‌ترین اطلاعات مورد نیاز، تعداد تقاضای روزانه اتوبوس است. با توجه به آنکه بیشتر ایستگاه‌های موجود در شهر مادرید دسترسی به بیش از یک خط را فراهم می‌کند و یا با فاصله کمی از یکدیگر قرار دارد و از نظر مشخصات محیط شهری یکسان است، تخمین تعداد مسافر بر اساس این مشخصات نتیجه ی یکسانی ایجاد می‌کند. بنابراین کل این ایستگاه‌ها به عنوان یک سیستم واحد در نظر گرفته شده و متغیر پیشنهادی، کل تقاضای اتوبوس است. با توجه به اثرگذاری تعداد خط در کل تقاضای اتوبوس، برای خنثی نمودن اثر تعداد خطوط اتوبوس، متغیر "تقاضای اتوبوس بر تعداد خطوط" در نظر گرفته شد. منبع اطلاعاتی هر دو متغیر، بر اساس اطلاعات گردآوری شده از یک روز کاری در بازه زمانی ۷ صبح الی ۱۰ شب در سال ۲۰۰۸ است. متغیرهای مستقل مدل نیز شامل مشخصات محیط شهری و مشخصات شبکه حمل‌ونقل است. در فرآیند مدلسازی، ابتدا با استفاده از مدل رگرسیون خطی دوگانه ارتباط بین هر متغیر و "تعداد کل تقاضا" و "تعداد کل تقاضا بر تعداد خطوط" بررسی شده، سپس مدل "تقاضای کل سفر با اتوبوس" و مدل "تعداد کل تقاضا بر تعداد خطوط" با استفاده از مدل رگرسیون خطی چندگانه برآورد شده و مشاهده شد، نتایج مدل کل تقاضای سفر بهتر بوده و متغیر مشخصات مکانی اهمیت زیادی در پیش‌بینی تقاضای اتوبوس دارد.

بر این اساس هدف اصلی مطالعه جاری ارائه روشی جهت پیش‌بینی تعداد مسافر خط اتوبوس در شهر تهران در افق کوتاه‌مدت (روزانه)، به منظور مدیریت هدفمند سرمایه‌های ملی، افزایش رضایت‌مندی مسافران سیستم اتوبوس‌رانی و برنامه‌ریزی عملیاتی خطوط مختلف اتوبوس است. اهمیت این موضوع در تهران به منظور افزایش کارایی و مطلوبیت سفر با توجه به آنکه در حال حاضر تعیین تعداد مسافر اتوبوس در هر خط با روش سنتی آماربرداری انجام شده و بسیار هزینه‌بر بوده، همچنین برنامه‌ریزی عملیاتی خطوط اتوبوسرانی به صورت تجربی توسط مدیر خط

انجام شده و هیچگونه برنامه‌ریزی مبتنی بر شرایط موجود برای آینده وجود ندارد، دوچندان است.

ساختار مقاله بدین شکل است که در بخش بعد به روش‌شناسی و در بخش سوم به تحلیل اطلاعات پرداخته می‌شود. در بخش چهارم نتایج پرداخت مدل و در بخش پایانی (پنجم) نیز، نتیجه‌گیری و پیشنهادات ارائه شده است.

۲. روش‌شناسی

به طور کلی اطلاعات موجود و ثبت شده توسط سیستم اتوبوسرانی شهر تهران، به دو دسته سیستم جمع‌آوری خودکار کرایه[۲۵] و سیستم موقعیت‌یاب خودکار وسیله نقلیه[۲۶] تقسیم می‌شود. جمع‌آوری خودکار کرایه یا بلیت الکترونیکی دانشی فنی و کاملا تخصصی است که در سیستم حمل‌ونقل عمومی شهر تهران شامل اتوبوس فیدر[۲۷] و سامانه اتوبوس تندرو[۲۸] راه‌اندازی شده است. مجموعه اطلاعات AFC، داده‌های ثبت‌شده حاصل از تراکنش‌های کارت‌بلیت‌های هوشمند برای پرداخت کرایه بوده و شامل اطلاعاتی همچون کد خط، تاریخ مصرف، زمان مصرف، کد دستگاه کارتخوان (شماره پلاک اتوبوس)، شماره سریال کارت‌بلیت، نوع کارت‌بلیت و کرایه‌ای که از کارت‌بلیت کسر شده، است (جدول ۱).

داده‌های ثبت شده توسط موقعیت‌یاب خودکار (AVL موجود در اتوبوس‌ها به منظور تعیین موقعیت جغرافیایی ناوگان اتوبوسرانی، موقعیت لحظه‌ای اتوبوس‌ها را با دقت قابل قبولی به مرکز اتوبوسرانی ارسال می‌کند و شامل اطلاعاتی همچون، تاریخ، شماره پلاک اتوبوس، شماره سریال GPS، کد ایستگاه ورودی، کد ایستگاه خروجی، زمان ورود اتوبوس به ایستگاه، زمان خروج اتوبوس از ایستگاه و زمان سفر طی شده بین دو ایستگاه است (جدول ۲).

به منظور برآورد تعداد مسافر خطوط اتوبوس، استفاده از اطلاعات AFC کافی بوده و با توجه به آنکه اطلاعات AVL بیانگر موقعیت لحظه‌ای اتوبوس (ایستگاه ورودی و خروجی) بوده و برای برآورد تعداد مسافر ایستگاه‌های اتوبوس در هر خط کاربرد دارد، در این مطالعه مورد استفاده قرار نگرفته است. در ادامه روش‌های مدلسازی تعداد مسافر خطوط اتوبوس ارائه شده است.

۲-۱ مدلسازی

به منظور برآورد تعداد مسافر خطوط اتوبوس از اطلاعات سیستم AFC برای هر خط و مدل‌های سری زمانی استفاده شده است. یک سری زمانی مجموعه منظمی از اطلاعات نقطه‌ای است که نوعا بر روی زمان‌های متوالی سنجیده می‌شود. از نظر ریاضی، سری زمانی به صورت یک مجموعه برداری $x_t, t = 0,1,2, \dots$ تعریف شده که t معرف زمان منقضی شده و متغیر x_t به عنوان یک متغیر تصادفی[۲۹] عمل می‌نماید. هدف اصلی این مدل‌ها، گردآوری دقیق و مطالعه قوی مشاهدات گذشته یک سری زمانی به منظور توسعه یک مدل مناسب برای توصیف ساختار ذاتی یک سری است. از این مدل‌ها برای تولید مقادیر آینده یک سری زمانی به منظور پیش‌بینی استفاده می‌شود. بطور کلی، مدل‌های مورد استفاده برای سری‌های زمانی به دو دسته مدل‌های تصادفی[۳۰] و شبکه عصبی مصنوعی[۳۱] تقسیم‌بندی می‌شود [Adhikari and Agrawal, 2013].

جدول ۱. نمونه‌ای از اطلاعات AFC

کد خط	تاریخ مصرف	زمان مصرف	شماره پلاک اتوبوس	شماره سریال کارت‌بلیت	نوع کارت‌بلیت	کرایه
٦٧٠١	٢٠-٠٢-١٣٩٤	١٦:٢٢	٦٤٢٣٤٦٥	٣٦٦٩٩٦٠٧٨٦	١٣٢	٥٦٤٠

جدول ۲. نمونه‌ای از اطلاعات AVL

تاریخ	شماره پلاک اتوبوس	سریال GPS	کد ایستگاه ورودی	کد ایستگاه خروجی	زمان ورود به ایستگاه	زمان خروج از ایستگاه	زمان سفر (ثانیه)
۱۳۹٤/۰۳/۱۸	٤۲۰۹٦٤۷۱۱	۹۸۹٥۹۰	٤	۳	۰۷:۱٤:٥۰	۰۷:۱۳:٥٤	٥٦

۲-۱-۱ مدل‌های تصادفی

یک مدل سری زمانی با توجه به نوع اطلاعات کنونی که می‌تواند تابعی خطی و غیرخطی از مشاهدات گذشته باشد، به صورت خطی یا غیرخطی تعریف می‌شود. به طور کلی مدل‌های سری زمانی شکل‌های مختلفی داشته و فرآیندهای تصادفی مختلفی را معرفی می‌نماید. دو مدل سری زمانی خطی که به طور گسترده در ادبیات مورد استفاده قرار می‌گیرند عبارتند از: مدل خودهمبسته (ار[۳۲]) و مدل میانگین متحرک (ما[۳۳]). از ترکیب این دو مدل، مدل‌های خود همبسته میانگین متحرک (ارما[۳٤]) و خود همبسته میانگین متحرک تلفیقی (اریما[۳٥]) ایجاد می‌شود. برای پیش‌بینی سری زمانی‌های فصلی نیز، مدل فصلی خودهمبسته میانگین متحرک تلفیقی (ساریما[۳٦]) استفاده می‌گردد [Adhikari and Agrawal, 2013].

برای تشخیص الگوی اطلاعات سری زمانی ابتدا باید تست ایستا انجام شود. بر این اساس می‌توان، با رسم نمودار خود همبستگی[۳۷]، نمودار خود همبستگی جزئی[۳۸] و شاخص دیکی- فولر[۳۹] الگوی اطلاعات از نظر تغییرات ثابت، غیر ثابت (نوسان) و دوره‌ای را تعیین نمود. در این مطالعه با توجه به توزیع ثابت و تغییرات نامنظم در سری زمانی مدل مناسب ساریما، انتخاب شد.

۲-۱-۲ مدل‌های شبکه عصبی

علاوه بر روش‌های تصادفی برای ساخت مدل‌های سری زمانی، مدل‌های شبکه عصبی مصنوعی روش دیگری برای پیش‌بینی سری‌های زمانی است که در سال‌های اخیر شهرت زیادی بدست آورده است. شبکه عصبی با استفاده از نرم‌افزار متلب[٤۰] به گونه‌ای طراحی می‌شود که قادر است؛ برآورد تعداد مسافر را از ابتدای شروع دوره در دسترس تا زمان t دریافت نموده و سپس با استفاده از این داده‌ها آموزش دیده؛ و برآورد تعداد مسافران در بازه زمانی tتا پایان دوره‌ای مورد نظر را پیش‌بینی کند.

یکی از ساده‌ترین و پرکاربردترین شبکه عصبی مصنوعی در پیش‌بینی مسائل، پرسپترون چندلایه[٤۱] است که از یک لایه پنهان برای پیش بردن شبکه استفاده می‌کند. این مدل با یک شبکه سه لایه‌ای ورودی، پنهان و خروجی مشخص شده است و ممکن است بیش از یک لایه پنهان وجود داشته باشد. گره‌ها در لایه‌های مختلف به عنوان المان‌های پردازشی شناخته می‌شوند [Adhikari and Agrawal, 2013]. خروجی مدل با استفاده از رابطه ۱ محاسبه می‌شود که در آن $y_{t-i}\ (i=1,2,\dots,p)$، ورودی و y_t خروجی مدل است. اعداد صحیح p و q به ترتیب تعداد ورودی و گره‌های پنهان است. ε_t خطای تصادفی، $\alpha_j (j=0,1,2,\dots,q)$ و $\beta_{ij} (i=0,1,2,\dots,p; j=0,1,2,\dots,q)$ وزن‌های ارتباطی و α_0 و β_{0j} ترم اریب[٤۲] است.

$$y_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j g\left(\beta_{0j} + \sum_{i=1}^{p} \beta_{ij} y_{t-i}\right) + \varepsilon_t \quad ,\forall\ t \qquad (۱)$$

معمولا تابع لوجستیک سیگموئید[٤۳] $g_x = \frac{1}{1+e^x}$ به عنوان تابع غیرخطی مؤثر بکار برده می‌شود. دیگر توابع مؤثر همچون خطی، تانژانت هیپربولیک[٤٤]، گوژین[٤٥] و غیره نیز استفاده می‌شود.

۲-۱-۳ روش‌های ارزیابی و اعتبارسنجی مدل

در این مطالعه برای ارزیابی مدل ساریما از آماره‌ی p-value، استفاده شده است. آماره‌ی p-value به منظور تعیین اهمیت هر یک از متغیرهای توضیحی مدل، استفاده می‌شود و بیانگر، میزان احتمال آنکه هیچ ارتباطی بین نمونه و جامعه وجود نداشته باشد، است. در بیشتر مطالعات، مقدار p-value ٥ درصد، به عنوان حداکثر خطای قابل قبول در نظر گرفته می‌شود، اما p-value ۱۰ درصد نیز قابل قبول است [Hensher, Rose and Greene, 2005].

به منظور مقایسه‌ی مدل‌های برآورد شده از آماره R^2[46] استفاده گردید، که نشان‌دهنده‌ی بهبود یا عدم بهبود مدل‌های ساخته شده نسبت به یکدیگر است. در واقع، R^2 بیانگر میزان تغییرات متغیر پاسخ (وابسته) است که توسط مدل قابل توضیح است و مقدار آن بین صفر و یک قرار می‌گیرد و بالا بودن آن نشان از توضیح‌دهندگی بیشتر مدل دارد.

با توجه به اهمیت اساسی پیش‌بینی سری‌های زمانی در بیشتر موقعیت‌های عملی، باید دقت لازم برای انتخاب مدل ویژه بکار گرفته شده و با استفاده از روش‌های مختلف، مدل‌های مختلف مقایسه و دقت پیش‌بینی هر یک تخمین زده شود. در این مطالعه معیار سنجش[47] مدل‌ها (مدل ساریما و شبکه عصبی پروسپترون چند لایه)، متوسط ریشه مربعات خطا[48] (رابطه ۲)، متوسط خطای مطلق[49] (رابطه ۳) و متوسط مطلق درصد خطا[50] (رابطه ٤) است که در آن، $e_t = y_t - f_t$ خطای پیش‌بینی، y_t مقدار واقعی، f_t مقدار پیش‌بینی شده و n اندازه مجموعه آزمایشی است [Adhikari and Agrawal, 2013].

$$\mathrm{MSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n} {e_t}^2} \qquad (۲)$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n} |e_t| \qquad (۳)$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{t=1}^{n} \left|\frac{e_t}{y_t}\right| \times 100 \qquad (٤)$$

۳. تحلیل اطلاعات موجود در سیستم اتوبوسرانی شهر تهران

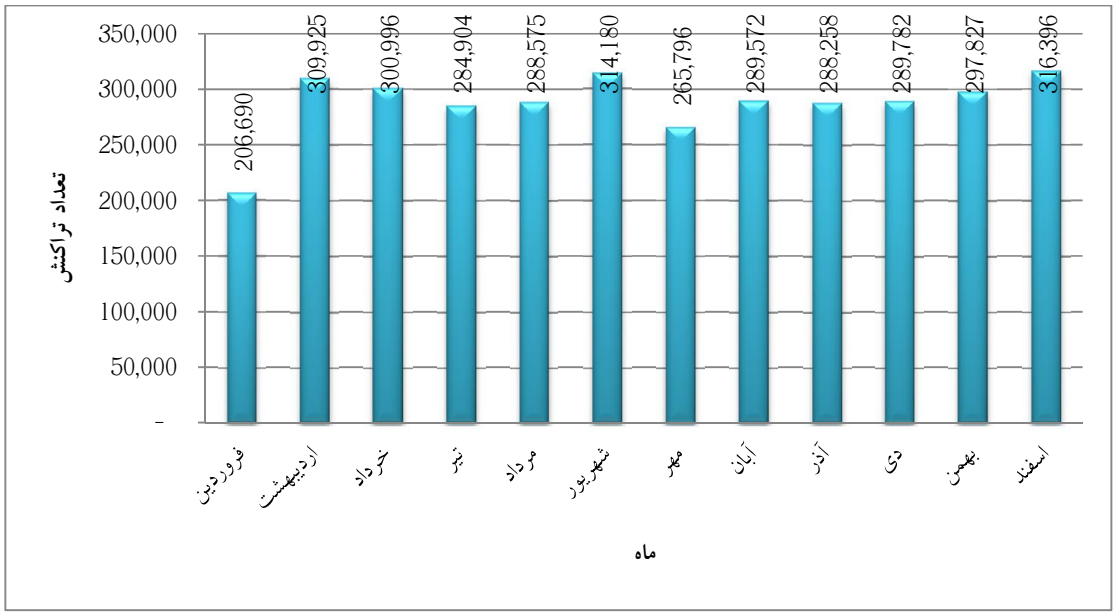
شبکه اتوبوسرانی شهر تهران شامل ۲۲۹ خط[51] بوده و به دو بخش اتوبوس‌های تندرو و فیدر تقسیم می‌شود. در حال حاضر در شهر تهران، ده خط اتوبوس تندرو وجود داشته و در تمامی خطوط مسافر هنگام ورود به ایستگاه کرایه را پرداخت می‌کند. اتوبوس‌های فیدر نیز که یکی از اهداف آنها رساندن مسافران به سامانه‌های خطوط تندرو است، شامل ۲۱۹ خط بوده و نحوه‌ی پرداخت کرایه هنگام پیاده شدن مسافر از اتوبوس است. در ادامه برای نمونه اطلاعات خط فیدر، پل کریمخان- میدان جمهوری (مسیر رفت با ۱۳ ایستگاه) و خط میدان جمهوری- پل کریمخان (مسیر برگشت با ١٤ ایستگاه)، مورد تحلیل قرار گرفته است.

۳-۱ تحلیل اطلاعات AFC

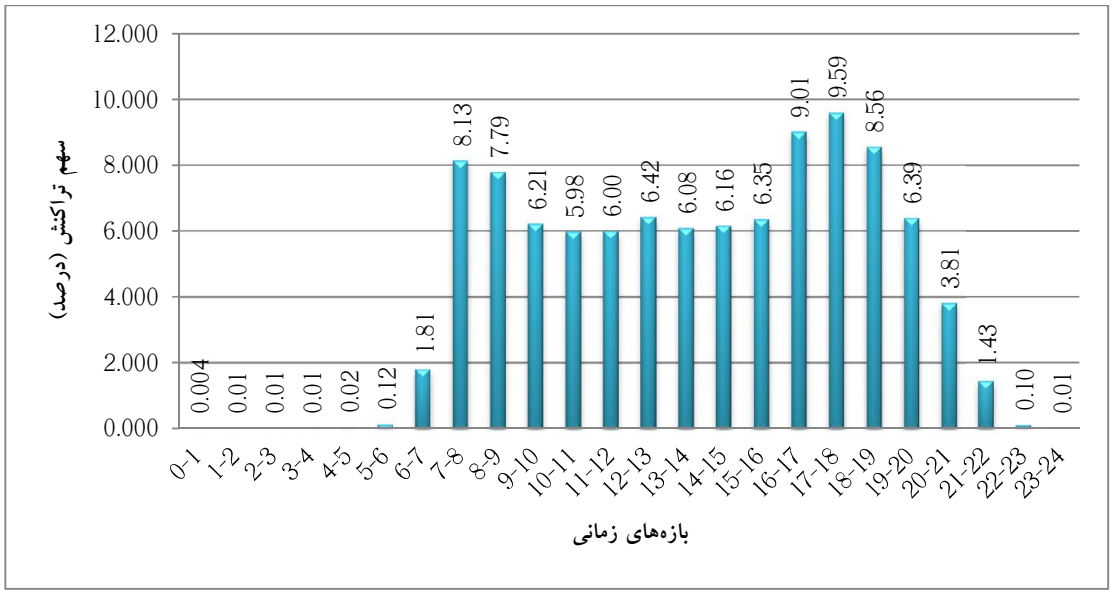
با توجه به تراکنش کارت‌بلیت‌ها از تاریخ ۱۳۹٤/۰۱/۰۱ الی ۱۳۹٤/۱۲/۲۹، در خط پل کریمخان- میدان جمهوری (رفت و برگشت) اطلاعات AFC مورد تحلیل قرار گرفت. بررسی تعداد تراکنش انجام شده در این خط نشان می‌دهد، در مجموع ۳٤٥۲۹۰۱ تراکنش در ۱۲ ماهه سال ۱۳۹٤ در این خط انجام شده، همچنین بررسی تعداد تراکنش به تفکیک ماه نشان می‌دهد (شکل ۱) بیشترین تراکنش به اسفند و کمترین تراکنش به فروردین ماه اختصاص دارد [Bus Tehran Information, 1394].

در شکل ۲ سهم بازه‌های زمانی مختلف از تراکنش سالیانه در خط پل کریمخان- میدان جمهوری ارائه شده و مشاهده می‌شود، زمان‌های دارای بیشترین تراکنش می‌تواند به دو بازه‌ی اوج صبح و اوج عصر تقسیم‌بندی شود. ساعت ۷ الی ۹ صبح بازه‌ی زمانی اوج صبح بوده که تعداد تراکنش در زمان ۷ الی ۸ به حداکثر می‌رسد. بازه‌ی زمانی ١٦ الی ۱۹ نیز بازه‌ی زمانی اوج عصر است که تعداد تراکنش در بازه‌ی زمانی ۱۷ الی ۱۸ به حداکثر تعداد خود رسیده است.

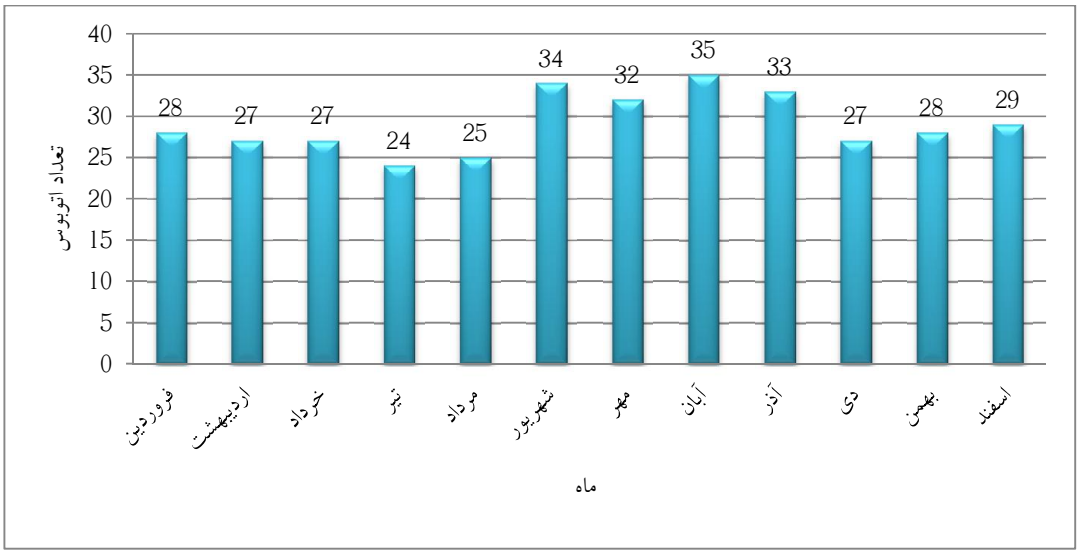
همانطور که بیان شد یکی از اطلاعات سیستم AFC کد اتوبوس یا همان پلاک اتوبوس است. تحلیل اطلاعات در سال ۱۳۹٤ نشان می‌دهد، ۷۷ دستگاه اتوبوس در خط پل کریمخان- میدان جمهوری مشغول به خدمت‌رسانی به مسافران بوده که این تعداد در ماه‌های مختلف سال متفاوت است و همانطور که در شکل ۳ مشاهده می‌شود، آبان ماه با ۳۵ اتوبوس بیشترین و تیر ماه با ٢٤ اتوبوس کمترین تعداد اتوبوس را به خود اختصاص دادند.



شکل ۱. تعداد تراکنش‌ها در خط پل کریمخان- میدان جمهوری به تفکیک ماه (۱۳۹٤)



شکل ۲. سهم بازه‌های زمانی مختلف از تراکنش سالیانه در خط پل کریمخان- میدان جمهوری (۱۳۹٤)



شکل ۳. تعداد اتوبوس‌های مشغول به فعالیت در خط پل کریمخان- میدان جمهوری به تفکیک ماه (۱۳۹٤)

٤. نتایج پرداخت و اعتبارسنجی مدل

در ادامه، نتایج مدل سری زمانی ساریما و روش شبکه عصبی پروسپترون چند لایه جهت برآورد تعداد مسافر خط پل کریمخان- میدان جمهوری ارائه شده و از داده‌های مربوط به مسافران جابجا شده از تاریخ ۹٤/۲/۱۸ تا ۹۵/۲/۱۸ جهت ساخت (آموزش مدل)، تست (آزمون) و اعتبارسنجی مدل استفاده شده است.

٤-۱ مدل ساریما (برآورد تعداد مسافر خط)

به منظور تعیین ایستایی مدل، آزمون دیکی- فولر تعمیم یافته انجام شد که نتایج آن در جدول ۳ ارائه شده و مشاهده می‌شود، سری زمانی ایستا است (چون آماره t دیکی- فولر از مقادیر بحرانی مربوط به سطوح معنی‌دار کوچکتر است). با توجه به ایستایی سری زمانی و تشخیص تغییرات نامنظم و فصلی آن، با استفاده از نرم افزار استتا[۵۲]، بهترین مدل از لحاظ دارا بودن کمترین مقادیر شاخص‌های خطا و بیشترین مقدار ضریب تعیین، مدل $Sarima(1,0,1)(1,0,1)^{14}$ تشخیص داده شد. لازم به ذکر است، در مدل‌های تصادفی (برای نمونه: ساریما) انتخاب داده برای مرحله آموزش و آزمون مدل، با توجه به اینکه داده‌ها به صورت سری زمانی بوده و باید داده پشت سر هم باشد، به صورت تصادفی نبوده و از بخشی از اطلاعات برای آموزش و از بخشی برای آزمون استفاده می‌شود. متغیرهای موثر در مدل، اطلاعات سری زمانی، به صورت روزانه جهت آموزش مدل از تاریخ ۹٤/۳/۱۸ تا ۹۵/۲/۱۸، متغیر ماه، روز هفته و تعطیلی بوده و خروجی مورد انتظار، پیش‌بینی تعداد مسافر خط در آن روز است. فرمول مدل پیشنهادی به صورت رابطه ۵ است که در آن:

- $y_{d(i)}$: تعداد مسافران برآورد شده خط پل کریمخان- میدان جمهوری در مدل به ازای هر روز i (i شمارنده روز)،
- m_j: متغیر دوتایی ماه (j شمارنده ماه)، w_k: متغیر دوتایی روز هفته (k شمارنده هفته) و h: متغیر دوتایی تعطیلی(۱=تعطیل رسمی)،
- a_d وایت نویز در سری زمانی با توزیع نرمال با میانگین صفر و انحراف معیار ۱۵۹۵ به ازای d درجه تلفیقی (۰=)،
- α_j: پارامتر ماه، β_k: پارامتر روز هفته، δ: پارامتر تعطیلی،
- φ_{d_i}: پارامتر خودهمبستگی به ازای d درجه تلفیقی (۰=) برای هر روز i (i شمارنده روز)، p درجه

خودهمبستگی، P درجه خودهمبستگی فصلی و s دوره تناوب فصلی، و

- θ_{d_i}: پارامتر میانگین متحرک به ازای d درجه تلفیقی (=۰) برای هر روز i (i شمارنده روز) ، q درجه میانگین متحرک، Q درجه میانگین متحرک فصلی و s دوره تناوب فصلی.

متغیرهای بکار رفته در مدل $Sarima(1,0,1)(1,0,1)^{14}$ و نتایج برآورد ضرایب مدل در جدول ٤ ارائه شده است.

جدول ۳. نتیجه آزمون دیکی فولر برای تعیین ایستایی مدل

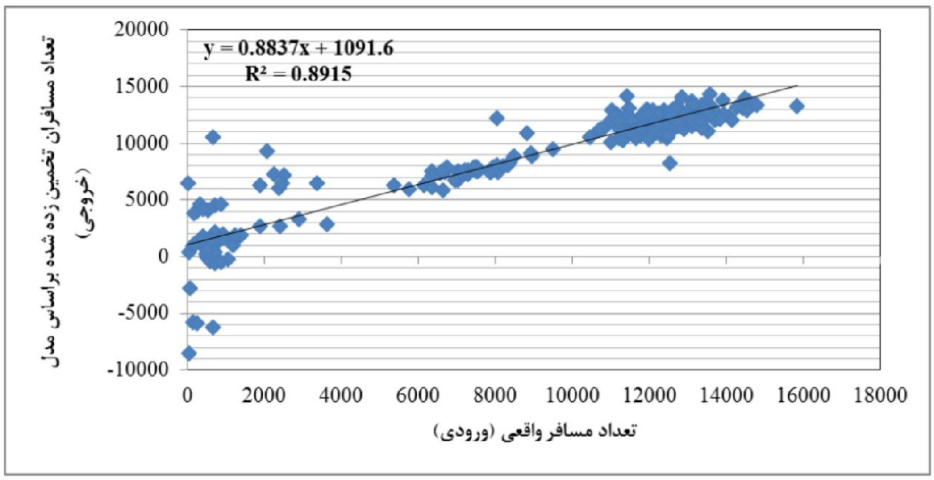
آزمون	آزمون دیکی فولر تعمیم یافته		آزمون مقادیر بحرانی	
آماره t	-۱۳/۸۹۵	-۳/٤٥۱	-۲/۸۷٦	-۲/٥۷۰
سطح معنی‌داری	-	سطح ۱٪	سطح ۵٪	سطح ۱۰٪

$$y_d(i) = \alpha_j m_j(i) + \beta_k w_k(i) + \delta h(i) + \sum_{l=1}^{p} \varphi_{d_l} y_d(i+l) + \sum_{n=s}^{P+s} \varphi_{d_P} y_d(i+n) + \sum_{t=P+s}^{P+p+s} \varphi_{d_{P+p}} y_d(i+t) + a_d(i) + \sum_{o=1}^{q} \theta_{d_o} a_d(i+o) + \sum_{r=s}^{Q+s} \theta_{d_Q} a_d(i+r) + \sum_{f=Q+s}^{Q+q+s} \theta_{d_{Q+q}} a_d(i+f) \quad (۵)$$

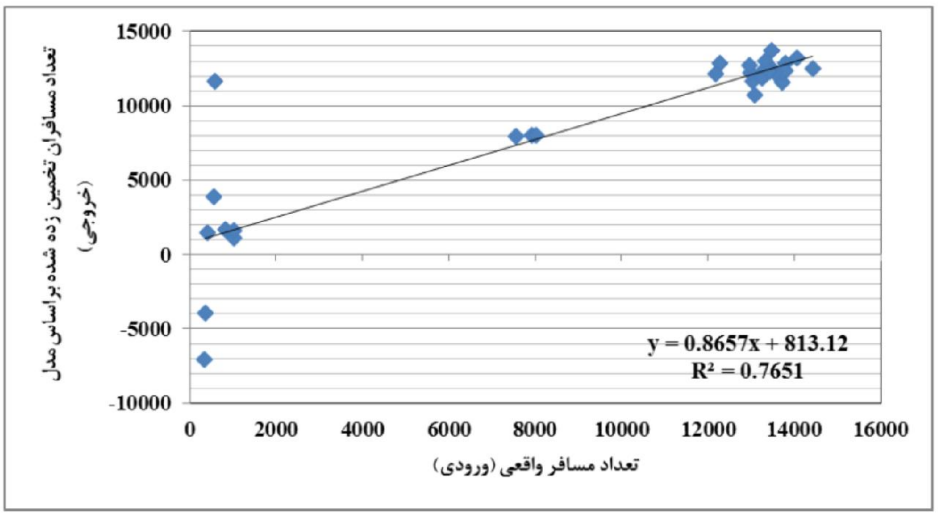
جدول ٤. نتایج برآورد ضرایب مدل $Sarima(1,0,1)(1,0,1)^{14}$

پارامتر	نام متغیر	تعریف	تخمین پارامتر	P-Value
C	ثابت	-	-۹۸۹/۸۷	۰/۰٥۷۰
$\propto_{2-6}$	متغیر ماه ۲ تا ماه ٦	متغیر دوتایی- ماه اردیبهشت تا شهریور =۱	۲۱۳٥/۷٤	۰/۰۰۰۰
$\propto_{7-11}$	متغیر ماه ۷ تا ۱۱	متغیر دوتایی- ماه مهر تا بهمن =۱	۲٥۸۰/۹۱	۰/۰۰۰۰
$\propto_{12}$	متغیر ماه ۱۲	متغیر دوتایی- ماه اسفند =۱	۳۷۹۷/٤۸	۰/۰۰۰۰
β_1	متغیر روز هفته ۱	متغیر دوتایی- روز شنبه =۱	۱۰٤۹٤/۹۰	۰/۰۰۰۰
β_2	متغیر روز هفته ۲	متغیر دوتایی- روز یکشنبه =۱	۱۰۷۱٤/۷۰	۰/۰۰۰۰
β_3	متغیر روز هفته ۳	متغیر دوتایی- روز دوشنبه =۱	۱۱۱۰۳/۱۹	۰/۰۰۰۰
β_4	متغیر روز هفته ٤	متغیر دوتایی- روز سه شنبه =۱	۱۰۸۷۱/٦۱	۰/۰۰۰۰
β_5	متغیر روز هفته ٥	متغیر دوتایی- روز چهارشنبه =۱	۱۰٤۸۰/۳۹	۰/۰۰۰۰
β_6	متغیر روز هفته ٦	متغیر دوتایی- روز پنجشنبه =۱	٦۰٥۸/۱٦	۰/۰۰۰۰
δ	متغیر تعطیلی	متغیر دوتایی- تعطیل رسمی =۱	-۷٥۹٤/۰۷	۰/۰۰۰۰
φ_{d_1}	پارامتر خودهمبستگی۱	درجه خود همبستگی=۱، درجه تلفیقی=۰	۰/٥۸۱۹	۰/۰۰۰۰
$\varphi_{d_{14}}$	پارامتر خودهمبستگی ۱٤	درجه خودهمبستگی فصلی=۱، با دوره تناوب= ۱٤، درجه تلفیقی فصلی=۰	-۰/۸٥۹۱	۰/۰۲۹۰
$\varphi_{d_{15}}$	پارامتر خودهمبستگی ۱٥	درجه خودهمبستگی=۱، درجه خودهمبستگی فصلی۱، با دوره تناوب=۱٤، درجه تلفیقی و درجه تلفیقی فصلی=۰	-۰/٤۹۹۹	۰/۰۰۰۰
θ_{d_1}	پارامتر میانگین متحرک ۱	درجه میانگین متحرک=۱، درجه تلفیقی=۰	-۰/۲٥۹۰	۰/۰۰٤۰
$\theta_{d_{14}}$	پارامتر میانگین متحرک ۱٤	درجه میانگین متحرک فصلی=۱، با دوره تناوب= ۱٤، درجه تلفیقی فصلی=۰	۰/۸۱٥۲	۰/۰٥۹۰
$\theta_{d_{15}}$	پارامتر میانگین متحرک ۱٥	درجه میانگین متحرک=۱، درجه میانگین متحرک فصلی=۱، با دوره تناوب= ۱٤، درجه تلفیقی و درجه تلفیقی فصلی=۰	-۰/۲۱۱۲	۰/۰۰۰۲

در شکل ٤، رگرسیون تعداد مسافر تخمین زده شده بر اساس مدل به تعداد مسافر واقعی در مرحله آموزش مدل، برای خط پل کریمخان- میدان جمهوری ارائه شده است. جهت ارزیابی و آزمون صحت‌سنجی مدل نیز از داده مربوط به تاریخ ۹٤/۲/۱۸ تا ۹٤/۳/۱۷ استفاده شده و نتایج رگرسیون تعداد مسافر تخمین زده شده براساس مدل به تعداد مسافر واقعی در مرحله آزمون مدل، در شکل ۵ ارائه شده است. معیارهای شاخص خطا جهت ارزیابی مدل ساریما برای دو مرحله آموزش و آزمون مدل نیز، در جدول ۵ ارائه شده است.



شکل ٤. رگرسیون تعداد مسافر تخمین زده شده خط به تعداد مسافر واقعی خط در مرحله آموزش مدل



شکل ۵. رگرسیون تعداد مسافر تخمین زده شده خط به تعداد مسافر واقعی خط در مرحله آزمون مدل

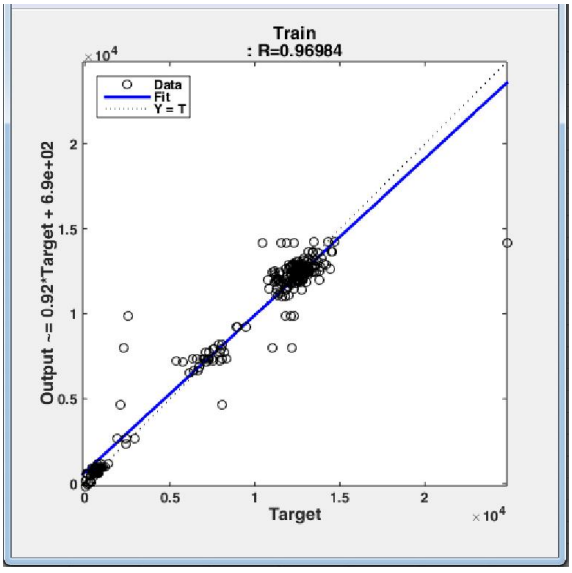
جدول ۵. مقادیر شاخص‌های خطا برای مدل تعداد مسافر خط، $Sarima(1,0,1)(1,0,1)^{14}$

شاخص خطا	تعداد داده	متوسط ریشه مربعات خطا	متوسط مطلق درصد خطا	متوسط خطای مطلق	R^2
آموزش مدل	۳۳۲	۱۵۹۰	۲٤	۹٤۲	۰/۸۹۱
آزمون مدل	۳۱	۲۷۰۱	۲۳	۱۵٤۵	۰/۷٦۵

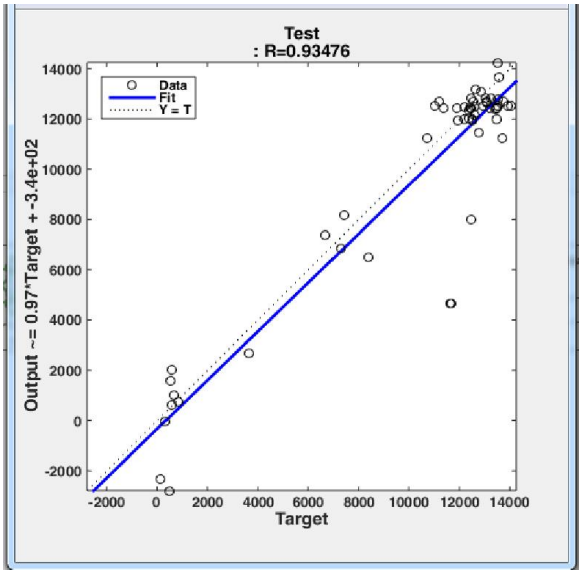
٤-٢ مدل پرسپترون چند لایه (مدل پیش‌بینی تعداد مسافر ایستگاه‌های اتوبوس)

برای پیاده‌سازی شبکه عصبی پروسپترون چند لایه از نرم‌افزار متلب استفاده گردید. برنامه به گونه‌ای نوشته شد که قادر است داده‌های دوره ابتدایی را دریافت کند، عمل نرمال‌سازی داده‌ها را انجام داده و در انتها با استفاده از این داده‌ها برآورد تعداد مسافر خطوط اتوبوس را پیش‌بینی نماید. برای پیدا کردن بهترین مدل پیش‌بینی، تعداد ورودی‌های مدل، تعداد لایه‌های میانی و تعداد گره‌های لایه-های میانی برای پیش‌بینی به برنامه داده شده و شاخص‌های عملکردی بر مبنای هر کدام محاسبه گردید. با تکرار آزمایش‌ها بر روی مقادیر مختلف ورودی، بهترین مدل، شبکه‌ای سه لایه‌ای با ۲۰ لایه ورودی، ۲۰ لایه پنهان (میانی) و ۱ لایه خروجی بدست آمد.

داده‌های مورد استفاده در ساخت این شبکه مربوط به تعداد مسافر جابجا شده در خط پل کریمخان- جمهوری در ۳٦۳ روز از تاریخ ۹٤/۲/۱۸ الی ۹۵/۲/۱۸ بوده، از ۷۰ درصد داده جهت آموزش شبکه، از ۱۵ درصد داده‌ها جهت ارزیابی شبکه و از ۱۵ درصد باقیمانده، جهت آزمون شبکه استفاده شده و انتخاب داده‌ها برای هر مرحله به صورت تصادفی است. ورودی شبکه شامل، تاریخ (مثال: ۹٤/۲/۱۸)، ماه (مثال: ۲)، هفته (مثال: جمعه: ۷)، تعطیل رسمی (مثال:۰) و تعداد مسافران استفاده کننده از خط پل کریمخان- جمهوری بر اساس روز است. نتایج رگرسیون تعداد مسافر خروجی به ورودی آموزش و آزمون مدل شبکه عصبی پروسپترون چند لایه در شکل ٦ ارائه شده است. معیارهای شاخص خطا جهت ارزیابی مدل پروسپترون چند لایه، برای دو مرحله آموزش و آزمون مدل در جدول ٦ نشان داده است.



آموزش



آزمون

شکل ٦. رگرسیون تعداد مسافر خط- خروجی به ورودی آموزش و آزمون شبکه عصبی پروسپترون چند لایه

جدول ٦. مقادیر شاخص‌های خطا برای مدل تعداد مسافر خط، مدل پروسپترون چند لایه

شاخص خطا	تعداد داده	متوسط ریشه مربعات خطا	متوسط مطلق درصد خطا	متوسط خطای مطلق	R^2
آموزش مدل	۲۵٤	۱۱۹۹	۲۰	٦۲٤	۰/۹۷۰
آزمون مدل	۵٤	۱۷۹۲	۳۷	۱۰۷۱	۰/۹۳۵

٤-٣ انتخاب مدل مناسب

در این مطالعه، مدل سری زمانی $Sarima(1,0,1)(1,0,1)^{14}$ و شبکه عصبی پروسپترون چند لایه، به عنوان بهترین مدل‌ها جهت برآورد تعداد مسافران خط پل کریمخان- میدان جمهوری انتخاب و پرداخت شد و در نهایت تعداد مسافر خط پل کریمخان- میدان جمهوری بر اساس هر دو مدل پیش‌بینی گردید. مقایسه نتایج شاخص‌های خطا برای هر دو مدل نشان می‌دهد، مدل پروسپترون چند لایه از لحاظ شاخص‌های خطا در پیش‌بینی تعداد مسافر خط از مدل ساریما برتر بوده و روش مناسب‌تری جهت برآورد تعداد مسافر خطوط اتوبوس با توجه به اطلاعات موجود است.

۵. نتیجه‌گیری و پیشنهادات

سیستم اتوبوسرانی یکی از بخش‌های مهم حمل‌ونقل بوده و عهده‌دار جابجایی تعداد زیادی از افراد و اقشار جامعه در سفرهای کاری و دیگر سفرهای روزمره است. یکی از مهم‌ترین راهکارها در حوزه حمل‌ونقل شهری انجام اقدامات لازم جهت توسعه و کارآمد نمودن این شیوه‌ی سفر است. بررسی‌ها نشان می‌دهد، برنامه‌ریزی برای استفاده بهینه از زیرساخت‌های اتوبوسرانی و بکارگیری راهکارهای افزایش کارایی این سیستم در دنیا، نیازمند وجود اطلاعات در زمینه زیرساخت و تقاضای سفر خطوط اتوبوس است. بر این اساس، در این مطالعه روش‌های مختلف برآورد تعداد مسافر خطوط اتوبوس بررسی شده و نتایج و پیشنهاداتی را به شرح زیر در برداشته است.

در این مطالعه، با استفاده از اطلاعات سیستم AFC (داده‌های ثبت‌شده حاصل از تراکنش‌های کارت‌بلیت‌های هوشمند برای پرداخت کرایه) موجود در سازمان اتوبوسرانی شهر تهران، مدل برآورد تعداد مسافر خطوط اتوبوس شهر تهران تهیه گردید. بدین منظور، ابتدا اطلاعات AFC مربوط به هر خط اتوبوس مرتب شد و بر اساس آن، مدل ساریما و شبکه عصبی (پرسپترون چند لایه) پرداخت گردید. نتایج مدلسازی نشان می‌دهد، مدل پروسپترون چند لایه از لحاظ شاخص‌های خطا همچون متوسط ریشه مربعات خطا، متوسط خطای مطلق و متوسط مطلق درصد خطا، در پیش‌بینی تعداد مسافر خط از مدل ساریما برتر بوده و روش مناسب‌تری جهت برآورد تعداد مسافر خطوط اتوبوس است.

خروجی مدل در این مطالعه، برآورد تعداد مسافر اتوبوس در هر خط برای روز مشخص است. بنابراین با توجه به آنکه تقاضای سفر با اتوبوس در روزهای مختلف سال (اسفند ماه، تعطیلات رسمی) و روزهای مختلف هفته (روزهای کاری و روزهای آخر هفته) متفاوت است، تعداد اتوبوس مورد نیاز برای هر خط در روزهای مختلف یکسان نیست. بنابراین با استفاده از تعداد مسافر پیش‌بینی شده برای هر خط می‌توان تعداد اتوبوس مورد نیاز برای هر خط در روزهای مختلف سال را تعیین نمود.

یکی دیگر از کاربردهای خروجی این مدل، تعیین معیار عملکرد اتوبوسرانی است. در بررسی عملکرد حمل‌ونقل عمومی دیدگاه‌های متفاوت و بعضا متناقضی وجود دارد. این دیدگاه‌ها در حالت کلی به سه دسته دیدگاه اداره‌کننده سیستم، دیدگاه کاربر و دیدگاه جامعه (محیط) تقسیم‌بندی می‌شود. اداره‌کنندگان به دنبال روشی هستند که به کمک آن بتوانند با حداقل استفاده از منابع در دسترس، سرویس مطلوبی تولید نماید. از طرف دیگر کاربران به دنبال سرویسی هستند که سرعت، راحتی، امنیت و قیمت مناسبی داشته باشد و محیط و جامعه هدف حداقل کردن اثرات منفی و استفاده از منابع را پیگیری می‌کنند. بررسی مطالعات مختلف نشان می‌دهد، یکی از شاخص‌هایی که در کارایی عملکردی خطوط اتوبوسرانی از دیدگاه اداره‌کننده مؤثر است، تعداد مسافرین سالیانه اتوبوس در هر خط است، که خروجی مدل حاصل از این مطالعه است.

با توجه به آنکه، پیاده‌سازی و بکارگیری راهکارهای افزایش کارایی سیستم اتوبوسرانی موجود در دنیا شامل، زمان‌بندی ناحیه‌ای[۵۳]، ارسال ناوگان خالی[۵۴]، سفرهای گردش کوتاه[۵۵] و پرش ایستگاهی[۵۶]، نیازمند اطلاعات دقیقی از تقاضای سفر ایستگاه‌ها در بازه‌های زمانی کوتاه‌مدت است و مطالعه جاری به پیش‌بینی تعداد مسافر اتوبوس در هر خط پرداخته، توصیه می‌شود برای مطالعات آتی، تعداد مسافر اتوبوس در ایستگاه‌های هر خط پیش‌بینی گردد.

٦. پی‌نوشت‌ها

1. Network Route Design
2. Timetable Development
3. Vehicle Scheduling
4. Crew Scheduling
5. Shenzhen
6. EWiews
7. Stationary test
8. Autoregressive Moving Average (ARMA)
9. Seasonal Autoregressive Integrated Moving Average (SARIMA)
10. ARIMA-GARCH
11. Interactive Multiple Model algorithm
12. Zilina
13. Slovak
14. SAS (Statistical Analysis System)
15. Autoregressive Integrated Moving Average (ARIMA)
16. Jinan
17. Autoregressive (AR)
18. Artificial neural network hybrid pattern (ANNHP)
19. GPS
20. Yantai
21. Time Varying Poisson Model
22. Weighted Time Varying Poisson Model
23. Weighted ensemble
24. Urban environment factor
25. Automated Fare Collection (AFC)
26. Automatic Vehicle Location (AVL)
27. Feeder bus
28. Bus rapid transit (BRT)
29. Random variable
30. Stochastic Models
31. Artificial Neural Networks
32. Autoregressive (AR)
33. Moving Average (MA)
34. Autoregressive Moving Average (ARMA)
35. Autoregressive Integrated Moving Average (ARIMA)
36. Seasonal Autoregressive Integrated Moving Average (SARIMA)
37. Autocorrelation function (ACF)
38. Partial autocorrelation function (PACF)
39. augmented Dickey-Fuller (ADF)
40. Matlab
41. Multi-layer perceptrons (MLPs)
42. Bias terms
43. Logistic sigmoid function
44. Hyperbolic tangent
45. Gaussian
46. R-squared
47. performance metrics
48. Root mean square error (RMSE)
49. The Mean Absolute Error (MAE)
50. The Mean Absolute Percentage Error (MAPE)

۵۱. بر اساس اطلاعات دریافتی ۲۳۰ خط در شهر تهران وجود دارد که وضعیت یکی از خطوط فیدر مشخص نبود و از تحلیل‌های انجام شده حذف گردید.

52. Stata
53. Zone Scheduling
54. Deadheading trip
55. Short turn trip
56. Skip Stop

۷. سپاسگزاری

از معاونت مطالعات و برنامه‌ریزی امور زیرساخت و طرح جامع، مرکز مطالعات حمل‌ونقل و برنامه‌ریزی شهر تهران، که حمایت مالی از این مطالعات را برعهده داشت و همکاران محترم آنن معاونت، صمیمانه سپاسگزاریم.

۸. مراجع

- شرکت مطالعات جامع حمل و نقل و ترافیک تهران (۱۳۸۳) "اطلاعات دریافتی از بانک اطلاعاتی مطالعات مبدأ- مقصد شهر تهران"، تهران: شرکت مطالعات جامع حمل‌ونقل و ترافیک تهران.
- شرکت واحد اتوبوسرانی (۱۳۹٤-۱۳۹۵) "اطلاعات دریافتی از شرکت واحد اتوبوسرانی تهران و حومه، واحد برنامه‌ریزی و مدیریت منابع"، تهران: شرکت واحد اتوبوسرانی

- Adhikari, R. and Agrawal, R. K. (2013) "An introductory study on time series modeling and forecasting", Lambert Academic Publishing (LAP), Cornell University, Germany.

- Ceder, A. (2007) "Public transit planning and operation", Taylor and Francis.

- Cyprich, O., Konecny, V. and Kilianova, K. (2013) "Short-term passenger demand forecasting using univariate time series theory", Promet, Traffic and Transportation, Vol. 25, No. 6, pp. 533-541.

- Hensher, D. A. Rose, J. M. Greene, W. H. (2005) "Applied choice analysis", Published in the United States of America by Cambridge University Press, New York.

- Maa, Z., Xing, J., Mesbah, M. and Ferreira, L. (2014) "Predicting short-term bus passenger demand using a pattern hybrid approach", Transportation Research Part C, Vol. 39, pp. 148–163.

- Pinedo, J. C. (2014) "Urban bus demand forecast at stop level: space Syntax and other built environment factors, evidence from Madrid", Procedia, Social and Behavioral Sciences, Vol. 160, pp. 205- 214.

- Xue, R. Sun, D. and Chen, S. (2015) "Short-term bus passenger demand prediction based on time series model and interactive multiple model approach", Discrete Dynamics in Nature and Society, No. 526, 11 pages.

- Zhoua, C., Daib, P., Wangc, F. and Zhanga, Z. (2016) "Predicting the passenger demand on bus services for mobile users", Pervasive and Mobile Computing, Vol. 25, pp. 48-66.

سید احسان سید ابریشمی، درجه کارشناسی در رشته مهندسی عمران را در سال ۱۳۸۲ از دانشگاه صنعتی شریف و درجه کارشناسی ارشد در رشته مهندسی عمران- مهندسی برنامه‌ریزی حمل‌ونقل را در سال ۱۳۸٤ از دانشگاه صنعتی شریف اخذ نمود. در سال ۱۳۸۹ موفق به کسب درجه دکتری در رشته مهندسی عمران- مهندسی برنامه‌ریزی حمل‌ونقل از دانشگاه صنعتی شریف گردید. زمینه‌های پژوهشی مورد علاقه ایشان، ارزیابی پروژه های حمل‌ونقل و مدلسازی تقاضا بوده و در حال حاضر عضو هیأت علمی با مرتبه استادیار در دانشگاه تربیت مدرس است.

وجیهه امینی، درجه کارشناسی در رشته مهندسی عمران را در سال ۱۳۹۰ از دانشگاه شاهد و درجه کارشناسی ارشد در رشته مهندسی عمران- مهندسی برنامه‌ریزی حمل‌ونقل را در سال ۱۳۹۲ از دانشگاه تربیت مدرس اخذ نمود. زمینه پژوهشی مورد علاقه ایشان، مدلسازی تقاضا است.

مریم ایران‌منش، درجه کارشناسی در رشته مهندسی عمران را در سال ۱۳۹۱ از دانشگاه بوعلی سینا همدان و درجه کارشناسی ارشد در رشته مهندسی عمران- مهندسی برنامه‌ریزی حمل‌ونقل را در سال ۱۳۹۵ از دانشگاه تربیت مدرس اخذ نمود. زمینه پژوهشی مورد علاقه ایشان، مدلسازی تقاضا است.

علی محدث دیلمی، درجه کارشناسی در رشته مهندسی برق- الکترونیک را در سال ۱۳۸۹ از دانشگاه علم و صنعت و درجه کارشناسی ارشد در رشته مهندسی برق گرایش مهندسی پزشکی را در سال ۱۳۹۱ از دانشگاه تهران اخذ نمود. در سال ۱۳۹٦ موفق به کسب درجه دکتری در رشته مهندسی برق، گرایش مهندسی پزشکی از دانشگاه تربیت مدرس گردید. علاقه پژوهشی ایشان شامل پردازش سیگنال و تصاویر دیجیتال، پردازش سیگنال های آرایه‌ای و مدلسازی است.