مجله علمی-پژوهشی
رایانش نرم و فن آوری اطلاعات
جلد ۴، شماره ۳، سال ۱۳۹۴

# پیش‌بینی زمان‌سفر در مسیرهای برون‌شهری با استفاده از تکنیک‌های داده‌کاوی مکانی مطالعه موردی: مسیر قائمشهر به بابل و ساری به قائمشهر

فردین مسلمی نجارکلائی[۱]، میثم عفتی[۲]، فرزین ناصری[۳]، محمدعلی رجبی[۴]
[۱] دانش آموخته کارشناسی ارشد مهندسی عمران - سیستم اطلاعات مکانی، گروه مهندسی سیستم های اطلاعات مکانی، دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته، کرمان، Fardin.moslemi@gmail.com
[۲] استادیار، گروه مهندسی عمران، دانشکده فنی، دانشگاه گیلان، رشت، Meysameffati@guilan.ac.ir
[۳]استادیار، گروه اکولوژی، پژوهشگاه علوم و تکنولوژی پیشرفته و علوم محیطی، دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته، کرمان، Fnnaseri@yahoo.com
[۴]استادیار، گروه مهندسی نقشه برداری، پردیس دانشکده‌های فنی، دانشگاه تهران، تهران، Marajabi@ut.ac.ir

**چکیده - برآورد زمان‌سفر در فرآیندهای برنامه‌ریزی سیستم‌های حمل‌ونقل، مورد توجه متخصصان حوزه‌های مختلف بوده که هر یک با بهره‌گیری از داده‌ها و ابزارهای تحلیلی خاص، سعی در مطالعه و برآورد زمان‌سفر داشته‌اند. روش‌های نوین و ترکیب ابزارهای موجود در سیستم‌های اطلاعات مکانی و تحلیل‌های داده‌کاوی، قابلیت‌های مختلفی را جهت استخراج الگوهای بدیع موجود در داده‌ها که بکارگیری آنها در برآورد زمان‌سفر فراهم آورده و جایگزین مناسبی برای روش‌های سنتی می‌باشند. این تحقیق از تکنیک‌های داده‌کاوی مکانی خوشه‌بندی K-means و همین‌طور روش شبکه عصبی پیش‌خور جهت برآورد زمان‌سفراستفاده می‌کند. در این تحقیق جهت برآورد زمان سفر از ۸ پارامتر استخراج شده از داده های GPS خودروهای عمومی و داده‌های هواشناسی و تصادفات در یک دوره زمانی ۳ ساله در استان مازندران (مسیر ساری- قائمشهر و بابل– قائمشهر) استفاده می‌شود. به‌منظور ارزیابی روش پیشنهادی، خروجی روش‌های خوشه‌بندی و شبکه‌عصبی در تعیین زمان سفر با داده‌های واقعی در مسیر مورد مطالعه مقایسه می‌شوند. نتایج حاکی از این است که روش شبکه‌عصبی مصنوعی با استفاده از روش مجذور میانگین مربع خطا با مقدار ۰/۰۰۱۵ و روش میانگین مربع خطا با مقدار ۰/۶۵ در مقایسه با روش‌های خوشه‌بندی نتایج رضایت بخشی دارد.**

**کلید واژه‌ها**– زمان سفر، سیستم اطلاعات مکانی، داده کاوی مکانی، شبکه‌عصبی مصنوعی پیش خور

## ۱. مقدمه

با افزایش جمعیت شهری و به تناسب آن افزایش ترافیک معابر، یکی از مهمترین دغدغه‌های مردم در مسیرهای برون-شهری زمان‌سفر در رسیدن به مقصد می‌باشد. برآورد زمان سفر در حوزه حمل‌ونقل جهت انتقال مردم از یک شهر به شهر دیگر و جابجایی اجناس و کالا بین دو شهر کاربرد دارد. اطلاعات مربوط به تاثیر شرایط آب‌وهوایی (آب و هوای معتدل، باران، برف، مه و یخبندان) و اطلاعات ترافیکی (تصادفات جاده‌ای و بازسازی راه‌ها) را می‌توان به‌عنوان پارامترهای تاثیرگذار بر روی جریان ترافیک دانست، این عوامل باعث تغییرات در سرعت خودرو شده و زمان رسیدن به مقصد را به‌صورت مستقیم تحت تاثیر قرار می‌دهند. بهره‌گیری از تکنولوژی‌های نوینی چون سامانه‌های هوشمند حمل‌ونقل و سامانه اطلاعات‌مکانی می‌تواند الگوهایی مناسبی را در زمینه هوشمندسازی سیستم حمل‌ونقل ارائه دهد. در سالیان اخیر با توسعه روش‌های جمع‌آوری داده و مواجهه با حجم عظیمی از داده‌های در دسترس، نیاز به روشهایی که به‌صورت خودکار قابلیت استخراج الگوهای مختلف از این منابع داده را داشته باشند، بیش از پیش احساس می‌شود. با پیشرفت علوم مهندسی، روش‌های داده‌کاوی به‌عنوان جایگزین روش‌های

سنتی قبلی جهت استخراج الگوهای سودمند، بدیع، قابل فهم و معتبر از داده‌ها مطرح شده‌اند. این روش‌ها با ترکیب روشهای یادگیری ماشین، هوش‌مصنوعی، پایگاه‌داده و غیره سعی در دستیابی به اطلاعات غنی نهفته شده در داده‌های ذخیره شده در پایگاه‌های داده را دارند [۱].

در این تحقیق با استفاده از روش‌های خوشه‌بندی و تکنیک شبکه‌عصبی مصنوعی و به کارگیری پارامترهای مناسب، به برآورد زمان سفر در مسیرهای ساری به قائمشهر و قائمشهر به بابل (استان مازندران- ایران) پرداخته شده است. بدین منظور ابتدا با استفاده از داده‌های موجود از منابع اطلاعاتی مختلف و با استفاده از روش‌های داده‌کاوی به آموزش سیستم پرداخته شده است. بعد از آموزش داده‌های تست وارد سیستم شده و زمان‌سفر با استفاده از الگوریتم‌های خوشه‌بندی اتصال‌منفرد و اتصال‌کامل، خوشه بندی‌افزایشی و تکنیک شبکه‌عصبی مصنوعی محاسبه می‌شوند. در نهایت نتایج حاصل از تکنیک‌های به‌کار رفته در برآورد زمان سفر را با داده‌های واقعی با استفاده از روش‌های مجذور میانگین مربع خطا و میانگین مربع خطای ارزیابی شده‌اند.

در بخش دوم و سوم مقاله به مرور تحقیقات پیشین و روش‌های پیش‌بینی زمان‌سفر و متغیرهای موثر بر زمان‌سفر پرداخته می‌شود. منطقه مورد مطالعه تحقیق در بخش چهارم آورده شده است. بخش پنجم به نحوه جمع‌آوری داده اختصاص داده شده است. بخش ششم به کدگذاری داده‌ها مربوط شده است. بخش هفتم به پیش‌پردازش داده‌ها پرداخته شده است. بخش نهم و دهم به ترتیب به برآورد زمان‌سفر با استفاده از مدل‌های خوشه‌بندی و شبکه عصبی مصنوعی پرداخته شده است. در پایان به بحث و ارزیابی نتایج پرداخته می‌شود.

## ۲. مروری بر تحقیقات پیشین

زمان‌سفر در مسیرهای برون‌شهری را می‌توان به دو روش مستقیم و غیرمستقیم به‌دست آورد. برآورد زمان‌سفر به روش مستقیم بر اساس فاصله زمانی عبور خودرو از دو نقطه ثابت تعیین می‌شوند. در روش‌های مستقیم با استفاده از فناوری‌های ردیابی وسایل نقلیه مانند خواندن پلاک خودروها از طریق دوربین‌ها، نرم‌افزارهای پردازش تصویر، سیستم‌های تعیین‌موقعیت‌جهانی (GPS) وسایل نقلیه خاص در دو نقطه از مسیر ردیابی می‌شوند و زمان واقعی سفر آنها در فاصله‌ای از مسیر در هر لحظه از روز اندازه‌گیری می‌شود. اما در روش غیر مستقیم بر اساس حجم، چگالی و سرعت خودرو می‌باشد که با استفاده از ابزارهای جمع‌آوری داده‌های ترافیکی همانند دوربین‌های ترافیکی و شناساننده‌های ترافیکی به‌دست آمده می‌شوند. به‌منظور برآورد زمان‌سفر با استفاده از روش‌های غیرمستقیم از روش‌های خاصی همانند روش‌های شبکه عصبی و روش‌های داده‌کاوی در بیشتر مطالعات گذشته استفاده شده است. برآورد زمان سفر در روش‌های غیرمستقیم به دلیل هزینه‌ها و مشکلات کمتری که نسبت به روش‌های مستقیم دارند مورد توجه بسیاری از مطالعات بوده است. تکنیک‌های پیش‌بینی زمان‌سفر بر اساس نحوه محاسبه زمان‌سفر، نحوه جمع‌آوری داده‌ها، مدل‌های پیش‌بینی قابل دسته‌بندی می‌باشند. از نظر نحوه محاسبه زمان‌سفر با تقسیم‌بندی کل مسیر به مسیرهای کوچک‌تر، زمان‌سفر هر یک از مسیرهای کوچک‌تر و متوالی محاسبه می‌گردد و در نهایت از جمع زمان‌های بدست آمده، زمان‌سفر کل مسیر محاسبه می‌گردد.

از مدل‌های پیش‌بینی گوناگونی در مطالعات گذشته جهت برآورد زمان‌سفر استفاده شده است. از جمله می‌توان به مدل‌های خطی و محاسبات آماری، خوشه‌بندی مکانی و شبکه‌عصبی‌مصنوعی اشاره داشت. امین ناصری و همکارانش (۱۳۸۷)، با استفاده از مدل رگرسیون خطی چندگانه و شبکه‌عصبی مصنوعی پیش‌خور به برآورد زمان سفر پرداخته‌اند. داده‌های مورد استفاده در این تحقیق به ده متغیر تقسیم گردیده و با استفاده از دوربین‌های قرار گرفته شده در سطح شهر جمع‌آوری شده‌اند. روش رگرسیون خطی چندگانه در فرآیندهای برنامه‌ریزی پویا نسبت به فرآیندهای برنامه‌ریزی استاتیک دارای دقت بالاتری می‌باشد. در روش شبکه عصبی مصنوعی، شبکه با استفاده از داده‌های گذشته طراحی و آموزش داده می‌شود و بر اساس داده های لحظه‌ای که از طریق حسگرهای ترافیکی و داده‌های سیستم‌های پیشرفته مدیریت ترافیک دریافت می‌شوند، تست می‌گردند. امین ناصری و همکارانش بیان نموده‌اند که برآورد زمان سفر با استفاده از شبکه‌عصبی مصنوعی دارای نتیجه بهتر و ضریب همبستگی بالاتری نسبت به مدل رگرسیون خطی چندگانه دارد [۴].

Wen و همکارانش (۲۰۰۲)، با استفاده از تکنیک شبکه‌عصبی به برآورد زمان‌سفر پرداخته‌اند. این تحقیق با استفاده از داده‌های جمع‌آوری شده توسط شناسنده‌های (دستگاه‌های جمع‌آوری اطلاعات ترافیکی از قبیل سرعت خودرو، حجم خودروهای عبور

کننده، فاصله بین خودروها و غیره) کنار جاده انجام گرفته و بر دو بخش استوار می باشد. در بخش اول این تحقیق با استفاده از روش فازی به یک الگوی مناسبی در زمینه برآورد زمان سفر دست پیدا می نمایند، و در بخش دوم تحقیق با استفاده از روش شبکه عصبی، مدل خوشه‌بندی شده را کالیبره کرده و نوسانات وجود داشته در میان داده‌های واقعی را از بین برده و در نهایت برای آنکه دقت خوشه‌ها را برآورد نمایند، از آنالیزهای حساسیت استفاده کرده‌اند [۵].

Nath و همکارانش (۲۰۱۰)، با استفاده از تکنیک داده‌کاوی مکانی و الگوریتم خوشه بندی K-Means به برآورد زمان سفر پرداخته‌اند. روش این الگوریتم این‌گونه است که داده‌های مشابه را در گروه‌های مجزا دسته‌بندی می‌نماید به‌طوریکه گروه‌های قرار گرفته‌شده در هر دسته هیچ مشابهتی با دسته‌های دیگر ندارند. داده‌های این تحقیق با استفاده از دستگاه GPS ای که درون خودروها می‌باشد، جمع آوری شده‌اند. در روش K-Means یک عدد یا یک عارضه به عنوان مرکز دسته شناخته می‌شود و اعداد دیگر یا عارضه‌های دیگر با استفاده از متد فاصله و یا مشابهت به آن مرکز دسته تعلق پیدا خواهند کرد. مزیت تکنیک خوشه‌بندی K-Means نسبت به تکنیک‌های دیگر خوشه‌بندی در این می‌باشد که در هر تکرار انجام شده برای بدست آوردن داده‌های مشابه در هر دسته، ممکن نیست جواب یکسان دهد و دسته‌ها مشابه با هم باشند. در نهایت با استفاده از فرمول های آماری زمان سفر مشخص می‌گردد [۶].

Chowdhury و همکارانش (۲۰۰۹)، با استفاده از دو تکنیک Moving Average و Chain Average به برآورد زمان سفر پرداخته‌اند. این دو تکنیک بر پایه یک سری فرمول‌های ریاضی ساده استوار می‌باشند. در این تحقیق برای برآورد زمان سفر از داده‌های پیشینی که توسط خودروها جمع آوری شده است، استفاده شده است و با استفاده از روش میانگین خطای نسبی مطلق دقت زمان سفر برآورد شده را ارزیابی می‌نمایند. روش Moving Average نسبت به روش Chain Average دارای دقت بهتری می‌باشد. از مزیت‌های این دو الگوریتم می توان اشاره به-سادگی، دقت بالا، هزینه‌های پایین و آنکه پیچیدگی سیستم به علت اجرای ساده هر دو روش پایین آمده و نوسانات یا نویزهایی که در مجموعه داده‌ها وجود داشته حذف می‌گردند [۷].

Chowdhury و همکارانش (۲۰۱۱)، به بررسی ضعف‌های تحقیقات انجام شده قبلی خویش پرداخته و الگوریتم‌های بهبود یافته K-Means، Moving Average و Chain Average را ارائه داده‌اند. از جمله ضعف‌هایی که الگوریتم‌های Moving Average وChain Average دارند می‌توان به این اشاره داشت که فضای محاسباتی این دو الگوریتم یک فضای بزرگی است یعنی مثلا اگر برای به‌دست آوردن زمان سفر n داده گذشته در دسترس باشد، ماتریس تولید شده برای برآورد زمان‌سفر یک ماتریس n×n خواهد بود (کاملا به داده‌ها وابستگی دارد). مشکل دوم این است که اگر جاده ای به چند بخش تقسیم شده باشد آنگاه نمی‌توان به طور مستقیم زمان‌سفر بخش آخری را محاسبه نمود و محاسبه آن کاملا به آرایه‌های ماتریس تشکیل شونده، بستگی دارد. مشکل دیگری که می‌توان در الگوریتم Moving Average به آن اشاره داشت این است که زمان برآورد شده یک بخش ممکن است به‌طور زیادی در برآورد زمان‌سفر بخش‌های دیگر دیده شود. الگوریتم K-Means هم دارای ضعف‌هایی می‌باشد که می‌توان اشاره به این داشت که عوارض به‌کار گرفته شده در یک فضای تک بعدی نسبت به هم دارای تجانس نمی‌باشند. در این تحقیق ضعف‌های این سه الگوریتم را بهبود داده و الگوریتم‌های توسعه یافته‌ای برای بهبود زمان سفر ارائه داده‌اند [۸].

با بررسی این مطالعات می‌توان دو عامل مهم را برای پیاده سازی تحقیق پیش رو نتیجه گرفت. موضوع اول راجع به پارامترهایی است که برای برآورد زمان‌سفر مناسب می‌باشند. در بسیاری از مطالعات مورد بررسی، تنها تعداد محدودی از پارامترها استفاده شده‌اند و یا پارامترهایی به کار رفته است که برخی از آنها مانند عرض خیابان و مسیر حرکت خودرو، با یکدیگر وابستگی دارند. بنابراین انتخاب داده‌های مناسب اولین قدم در برآورد زمان‌سفر می‌باشد. لذا در انتخاب داده‌ها سعی بر آن شده است که نسبت به یکدیگر دارای همبستگی کمی باشند و بیشترین میزان تاثیر در مسیرهای برون شهری بر زمان‌سفر داشته باشند. موضوع دوم در ارتباط با روش مدل‌سازی پارامترها می‌باشد. طبق مطالعات انجام شده روش شبکه عصبی مصنوعی در برآورد زمان سفر نسبت به روش‌های خطی از میانگین خطای نسبی کمتر، ضریب همبستگی بالاتر و قابلیت وارد کردن داده‌های حال برخوردار می‌باشد. روش خوشه‌بندی K-Means نسبت به روش‌های Chain Average و Moving Average دارای پیچیدگی محاسباتی کمتر و از دقت بالاتری برخوردار می‌باشد و می‌توان پارامترهای مختلف تاثیرگذار در زمان‌سفر را با استفاده از مقیاس فاصله در خوشه‌های مجزا طبقه‌بندی کرد و زمان‌سفر مربوط به هر خوشه، که داده‌های آن با یکدیگر مشابه می‌باشند و نسبت به خوشه‌های دیگر متفاوت هستند را برآورد کرد. لذا به علت دلایل اشاره شده از روش‌های شبکه‌عصبی‌مصنوعی و خوشه بندی K-Means استفاده شده است.

در این تحقیق تاثیر همزمان پارامترهای برون جاده‌ای (شرایط آب‌وهوایی مختلف و شرایط تصادفات جاده‌ای و زمان‌های مختلف سفر) بر زمان‌سفر به‌دست آمده‌اند. پارامترهای مورد استفاده از وابستگی کمی نسبت به یکدیگر برخوردارند. قابلیت برآورد زمان‌سفر با استفاده از داده‌های حال در روش‌های مورد استفاده وجود دارد.

## ۳. متغیرهای موثر بر اندازه‌گیری زمان‌سفر

پارامترهای زیادی در برآورد زمان‌سفر تاثیر دارند که شامل داده‌های حجم ترافیکی، شرایط آب‌وهوایی، داده‌های Gps خودروهای عمومی (داده‌های Gps خودروهای عمومی شامل موقعیت خودرو، سرعت متوسط حرکت خودرو، تاریخ، روز، زمان شروع و پایان حرکت خودرو و تاریخ می‌باشد) و شرایط مکانی سفر می‌باشند.

### ۳-۱ داده‌های ترافیک

از میان عوامل‌های تاثیرگذار بر زمان‌سفر، شرایط ترافیکی مهم‌ترین بخش تاثیرگذار بر زمان‌سفر است. سرعت وسایل نقلیه در یک مسیر، در کنار متغیرهای ترافیکی دیگر مانند تردد، حجم و ازدحام و یا باز بودن مسیر را نشان می‌دهد. تعاریف و روابط متعددی برای محاسبه هر یک از متغیرهای ترافیکی ارایه شده است که بخشی از آنها طبق نشریه ۱۱۶ طرح هندسی راه عبارت‌اند از:

***تردد مسیر:*** تردد عبارت است از تعداد خودروی عبوری از یک نقطه مسیر در طول زمانی ثابت. چنانچه (M) تعداد خودروی عبوری از یک نقطه مسیر در فاصله زمانی (T) باشد، میزان متغیر تردد که با (q) نشان داده می‌شود طبق رابطه ۱ بیان می‌گردد.

$$q = \frac{M}{T} \qquad (۱)$$

***فاصله عبور:*** متوسط فاصله زمانی یا مسافتی بین دو وسیله نقلیه متوالی در حال حرکت در یک مسیر، فاصله عبور گفته می‌شود.

***سرعت خودرو:*** بیانگر سرعتی است که یک خودرو بر حسب کیلومتر بر ساعت دارا می‌باشد.

***زمان اوج ترافیک:*** بیانگر زمانی است که تعداد وسایل نقلیه در یک بازه زمانی به حداکثر خود در یک مسیر مدنظر می‌رسند.

***زمان بدون اوج ترافیک:*** بیانگر زمانی است که تعداد وسایل نقلیه در یک بازه زمانی به حداقل خود در یک مسیر مد نظر می‌رسد.

### ۳-۲ داده‌های آب‌وهواشناسی

حوادث آب‌وهوایی مثل بارش، مه، بادهایی با شدت وزش بد، میزان بارش زیاد (تگرگ) و برف سبب کاهش ظرفیت راه می‌شوند. این حوادث می‌توانند موجب کاهش سرعت حرکت در خودرو، کاهش حجم تردد خودرو و تاخیر زیاد در رسیدن از شهر مبدا به شهر مقصد گردند. این اطلاعات از ایستگاه‌های هواشناسی مسیر مورد مطالعه جمع آوری می‌شوند. با آگاه بودن از شرایط آب‌وهوایی مسیر مورد مطالعه می‌توان تاثیر شرایط آب‌وهوایی نامساعد را بر روی زمان سفر به‌دست آورد.

### ۳-۳ داده‌های تصادفات جاده‌ای

تصادفات جاده‌ای یکی از عواملی می‌باشد که منجر به مسدود شدن راه و بالا رفتن حجم ترافیک می‌شود و نتیجه آن سبب تاخیر در زمان رسیدن از شهر مبداء به شهر مقصد می‌باشد. می‌توان اطلاعات شرایط تصادفات جاده‌ای را توسط پلیس راه‌ها جمع‌آوری نمود. تصادفات در یک مکان و یک زمان خاصی صورت می‌گیرد. با ترکیب تاریخ وقوع تصادف، مکان وقوع تصادف و زمان وقوع تصادف با مسیر، تاریخ و زمان حرکت خودرو می‌توان اتفاق و یا عدم اتفاق افتادن تصادف در مسیرحرکت خودرو را متوجه شد. اگر تصادفی در مسیر حرکت خودرو اتفاق افتاده باشد می‌توان میزان تاخیری که در رسیدن خودرو به شهر مقصد به‌وجود آمده است را نسبت به شرایط ایده‌آل (عدم اتفاق افتادن تصادف) به‌دست آورد. با استفاده از روش‌های خوشه‌بندی با به‌دست آمدن خوشه‌هایی که دارای شرایط تصادف در مسیر حرکت خودرو می‌باشند، زمان‌سفر با استفاده از حاصل‌ضرب طول راه در میانگین سرعت‌های متوسط قرار گرفته شده در خوشه برآورد می‌شود.

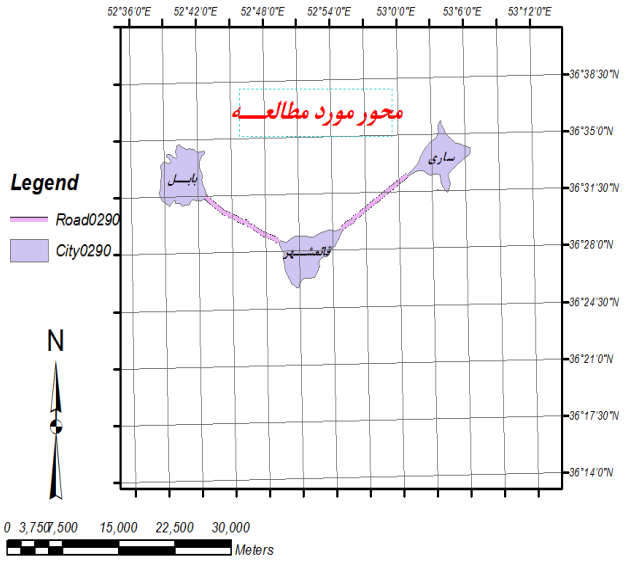
### ۳-۴ شرایط مکانی سفر

منظور از شرایط مکانی سفر (شرایط تسهیلات ثابت)، ویژگی‌های خیابان‌ها و مسیرهایی است که وسیله نقلیه در آنها تردد می‌کند. تعداد چراغ‌های راهنمایی در طول مسیر، تعداد تقاطع‌های مسیر،

اختلاف ارتفاع مقصد از مبدأ و نوع روسازی مسیر و ... از جمله متغیرهایی هستند که تأثیر ویژگی‌های تسهیلات ثابت را بر زمان سفر اندازه می گیرند. این دسته از متغیرها، بالقوه می‌توانند به‌طور مستقیم یا غیر مستقیم بر زمان سفر وسایل نقلیه تأثیر بگذارند. به‌عنوان مثال روسازی نامناسب خیابان‌ها به‌طور غیرمستقیم با تأثیر بر سرعت وسایل نقلیه عبوری در مسیر باعث افزایش زمان سفر می‌شود و یا تعداد چراغهای راهنمایی طول مسیر به‌طور مستقیم بر زمان سفر تأثیر دارد.

## ۴. منطقه مورد مطالعه

منطقه مورد مطالعه مسیرهای برون شهری استان مازندران محور ساری به بابل می‌باشد. محور ساری به بابل به علت محور متصل‌کننده استان مازندران به استان تهران به عنوان یکی از پرترددترین محورهای ایران می‌باشد. این مسیر با طول ۳۹۷۹۴/۴۱ متر و مختصات نقطه شروع از شهرستان بابل در سیستم جهانی یونیورسال ترانسورس مرکاتور برابر با $E = 650388/38m,\ N = 4044770/7m,\ h = 132m$ متر و نقطه پایانی مسیر در شهرستان ساری در سیستم جهانی ترانسورس مرکاتور دارای مختصات $E = 684231.25m,\ N = 44048683.50m,\ h = 76m$ متر می‌باشد. شکل ۱ نمای کلی محور مورد مطالعه که ما بین سه شهر ساری قائمشهر و بابل قرار دارد را نشان می‌دهد.




شکل۱. نمای کلی محور مورد مطالعه

| | نوع داده | منبع جمع آوری داده |
|---|---|---|
| ۱ | داده‌های GPS خودروهای عمومی (سرعت‌متوسط خودرو، بازه‌زمانی حرکت‌خودرو، مسیر حرکت‌خودرو، مدت‌زمان حرکت خودرو، روز حرکت خودرو و تاریخ حرکت خودرو) | ستاد فرماندهی پلیس راههای استان مازندران |
| ۲ | داده‌های تصادف | ستاد فرماندهی پلیس راههای استان مازندران |
| ۳ | اطلاعات حجم ترافیکی (زمان‌اوج و زمان‌بدون اوج) | ستاد فرماندهی نیروی انتظامی استان مازندران |
| ۴ | داده های هواشناسی (آب‌وهوای آفتابی و بارانی و باد) | ایستگاه هواشناسی قائمشهر |

شکل۲. داده‌های مورد استفاده در تحقیق به همراه منبع جمع‌آوری آنها

## ۵. جمع آوری داده

به‌منظور برآورد زمان‌سفر، داده‌هایی را به عنوان ورودی در مدل گرفته و خروجی (زمان‌سفر) برآورد می‌شود. داده‌های مورد نیاز جهت انجام کار عملی تحقیق حاضر شامل داده‌های لازم جهت برآورد زمان‌سفر می‌باشند. لذا در شکل ۱ لیستی از آنها ملاحظه می‌شود و به معرفی آنها پرداخته می‌شود.

همچنین از نرم افزارهایی برای پردازش داده‌ها استفاده گردید. نرم افزارهای Microsoft SQL Server 2008 و Microsoft Office Excel 2007 برای ذخیره و جمع‌آوری داده‌ها، نرم افزار Microsoft Visual Studio 2010 و Matlab به‌منظور پیش‌پردازش و پردازش داده‌ها مورد استفاده قرار گرفت.

زمان اوج به‌عنوان یکی از داده‌های حجم ترافیک باعث کاهش سرعت حرکت خودرو و در نتیجه تاخیر در زمان رسیدن به مقصد می‌گردد و می‌تواند یکی از عوامل موثر در برآورد زمان‌سفر باشد. به‌علت آنکه در زمان اوج ترافیک، میزان تردد عبور و مرور خودرو به حداکثر مقدار می‌رسد لذا زمان‌سفر در مسیر حرکت خودرو افزایش می‌یابد. طبق نظر کارشناسان برای منطقه مورد مطالعه از نظر زمان اوج ۲ بازه زمانی در نظر گرفته شده است. زمان اوج حرکت و زمان بدون اوج حرکت مسیر مورد مطالعه شهرستان‌های استان مازندران طبق جدول ۱ می‌باشد.

جدول۱. جدول زمان اوج و بدون اوج

| گروه زمانی | نوع زمانی | بازه زمانی | گروه زمانی | نوع زمانی | بازه زمانی |
|---|---|---|---|---|---|
| ۴ | اوج حرکت | ۱۸ - ۱۶ | ۱ | اوج حرکت | ۹/۳۰- ۶/۳۰ |
| ۵ | بدون اوج حرکت | ۲۴ - ۱۸ | ۲ | بدون اوج حرکت | ۱۳- ۹/۳۰ |
| ۶ | بدون اوج حرکت | ۶/۳۰- ۰۰ | ۳ | بدون اوج حرکت | ۱۶ - ۱۳ |

به‌منظور استخراج پارامترهای مورد نیاز برآورد زمان‌سفر از داده‌های GPS خودروهای عمومی (زمان شروع حرکت خودرو، زمان پایان حرکت خودرو، سرعت حرکت خودرو، تاریخ حرکت خودرو) استخراج می‌شود. نحوه استخراج پارامترهای مورد نیاز این‌گونه است که پس از باز کردن فایل‌های مربوط به GPS خودروها در نرم‌افزارردیاب (تنها نرم‌افزاری است که می‌تواند اطلاعات درون GPS خودروهای عمومی را در ایران باز نماید.) تمام تاریخ‌هایی که خودرو در راه‌های مختلف کشور حرکت کرده است، نشان داده می‌شود. پس از مشاهده تاریخ حرکت خودرو و باز نمودن فایل مربوط به هر تاریخ می‌توان داده‌های مدت‌زمان حرکت خودرو، سرعت حرکت خودرو، تاریخ و زمان (ساعتی که خودرو حرکت کرده است) را استخراج کرد.

چون سیستم GPS خودرو در هر یک دقیقه سرعت خودرو را ثبت می‌نماید لذا در مسیر حرکت خودرو سرعت‌های مختلفی وجود دارد. با به‌دست آمدن سرعت‌های مختلف خودرو با میانگین‌گیری سرعت‌ها یک سرعت متوسط به منظور مبنا گرفتن سرعت حرکت خودرو در کل مسیر تعریف می‌شود. واحد سرعت متوسط کیلومتر بر ساعت می‌باشد.

پارامتر تاریخ و ساعت حرکت خودرو را نیز با استفاده از نرم‌افزار ردیاب استخراج می‌شود. با ترکیب پارامتر تاریخ و ساعت و مکان حرکت خودرو می‌توان وقوع تصادف و شرایط مختلف آب‌وهوایی را در مسیر حرکت خودرو متوجه شد و میزان تاثیر پارامتر تصادف و شرایط آب‌وهوایی مختلف را در زمان‌سفر به‌دست آورد.

اطلاعات مربوط به تصادفات جاده‌ای از طریق پلیس راه‌های استان مازندران جمع آوری می‌شوند. اطلاعات تصادفات جاده‌ای در قالب ستون‌هایی از جمله تاریخ تصادف، محل حادثه، مقصر و ساعت تصادف تعریف می‌شوند و تحت فایل Excle ذخیره می‌گردند. از اطلاعات تصادفات جاده‌ای نیازمند به پارامترهای ساعت، تاریخ و مکان تصادف خواهیم بود. با استفاده از پارامتر-های اشاره شده می‌توان پی به وقوع تصادف در مسیر حرکت خودرو برد.

محور ساری به بابل در میان سه شهر ساری قائمشهر بابل قرار گرفته شده است که در این مسیر۱ ایستگاه سینوپنتیک به نام ایستگاه سینوپتیک قراخیل که محدوده این مسیر را تحت پوشش قرار می‌دهد، برای ثبت اطلاعات هواشناسی در نظر گرفته شده است.

## ۶. کدگذاری داده‌ها

پس از جمع‌آوری پارامترهای مورد نیاز تحقیق که به‌ترتیب عبارتند از سرعت متوسط حرکت خودرو، مدت زمانی که خودرو در طول مسیر حرکت کرده است، تاریخ حرکت خودرو، بازه زمانی حرکت‌خودرو، شرایط آب‌وهوایی، شرایط تصادفات‌جاده‌ای، روزهای هفته و یکپارچه‌سازی پارامترها، فرآیند کدگذاری داده‌ها انجام می‌شود. به‌منظور سرعت بخشیدن و تسهیل در فرآیند تجزیه و تحلیل داده‌ها، مسیر حرکت خودرو، بازه‌های زمانی حرکت‌خودرو، شرایط آب‌وهوایی و شرایط تصادفات کدگذاری می‌شوند. زمان حرکت خودرو به زمان‌های اوج و غیراوج صبح،ظهر، اوج و غیراوج بعدازظهر و شب کدگذاری می‌شوند. از ساعت ۶/۳۰ تا ۹/۳۰ صبح کد ۱، ۹/۳۰ تا ۱۳ کد ۲، ۱۳ تا ۱۶ کد ۳، ۱۶ تا ۱۸ کد ۴، ۱۸ تا ۲۴ کد ۵ و ۰/۰ تا ۶/۳۰ کد ۶ اختصاص داده شده است. به مسیر مابین دو شهر مبداء و مقصد یک کدی اختصاص داده می‌شود. در این تحقیق به مسیر مابین شهر ساری و قائمشهر کد ۱ و به مسیر حرکتی مابین شهر بابل و قائمشهر کد ۲ را اختصاص داده شده است. کدگذاری شرایط تصادف در طول محور به اینگونه می‌باشد که اگر در بازه زمانی حرکت خودرو تصادفی اتفاق افتاده باشد با کد ۱ مشخص می‌گردد و اگر هیچ تصادفی رخ نداده باشد با کد ۲ مشخص می‌گردد. کدگذاری شرایط آب‌وهوایی طبق جدول ۲ در نظر گرفته شده است. جدول ۳ بیانگر جدول داده‌های آماده‌سازی شده در نرم‌افزار می‌باشد.

جدول ۲. کدگذاری شرایط آب و هوایی

| شماره | شرایط آب و هوایی در #C | شرایط آب و هوایی در مطلب |
|---|---|---|
| ۱ | آفتابی | ۱۰ |
| ۲ | بارندگی | ۱۱ |
| ۳ | باد | ۱۲ |

جدول ۳. نمونه ای از داده‌های آماده سازی شده

| آب‌وهوایی | کد تصادف | گروه زمانی | کد شهر | سرعت Km/h | زمان (دقیقه) | مقصد | مبدا |
|---|---|---|---|---|---|---|---|
| آفتابی | ۰ | ۳ | ۱ | ۸۹/۲۵ | ۱۲ | قائمشهر | ساری |
| بارندگی | ۰ | ۳ | ۲ | ۵۷/۱۹ | ۱۶ | قائمشهر | بابل |
| آفتابی | ۰ | ۴ | ۱ | ۸۲/۰۷ | ۱۳ | قائمشهر | ساری |
| آفتابی | ۱ | ۶ | ۲ | ۵۲ | ۲۰ | قائمشهر | بابل |
| بارندگی | ۰ | ۶ | ۱ | ۴۹/۷۲ | ۱۸ | قائمشهر | بابل |
| آفتابی | ۰ | ۵ | ۲ | ۵۰/۵۶ | ۲۳ | قائمشهر | بابل |
| آفتابی | ۰ | ۵ | ۱ | ۶۹/۳۳ | ۱۵ | قائمشهر | ساری |
| بارندگی | ۱ | ۶ | ۱ | ۳۱/۱۲ | ۳۷ | قائمشهر | ساری |

## ۷. پیش‌پردازش داده‌ها

بعد از گردآوری داده‌ها بایستی این اطمینان حاصل شود که داده‌ها برای ارزیابی و بدست آوردن یک نتیجه درست مناسب می‌باشند. بنابراین پیش پردازش داده‌ها مهم و ضروری می‌باشد. فرآیند پیش‌پردازش دده ها به‌طور کلی شامل دو مرحله می‌باشد:

- آشکارسازی و حذف داده‌هایی که دارای خطا هستند.
- مقیاس‌بندی و انتخاب

به‌منظور تفکیک شرایط‌های مختلف آب‌وهوایی، زمانی و تصادفات و پیاده‌سازی شبکه عصبی مصنوعی و خوشه‌بندی K-means باید انتخاب در میان داده‌ها انجام گردد. انتخاب اولیه منجر به این می‌شود که تمامی شرایط مشابه از جمله گروه‌های زمانی، شرایط مختلف آب‌وهوایی و تصادفات در گروهای مشابه قرار گیرند. گروه‌ها تنها در سرعت متوسط دارای تمایز می‌باشند. در پردازش نهایی سرعت متوسط گروه‌ها نسبت به یک حد آستانه، مشابه یکدیگر می‌شوند. به‌عنوان مثال تمامی خودروهایی که در بازه زمانی ۶ صبح تا ۹ صبح در مسیر ساری به قائمشهر و شرایط آب‌وهوایی آفتابی بدون هیچ گونه تصادف و در یک روز عادی در طول مسیر جاده حرکت نموده‌اند، دسته‌بندی می‌شوند. این گروه‌بندی سبب می‌گردد که پردازش‌نهایی با سرعت بالا و زمان کمتری انجام گیرد. پس از آماده‌سازی داده‌ها برای انجام فرآیند دسته‌بندی، داده‌ها وارد پایگاه داده می‌گردند. در این تحقیق از پایگاه داده Microsoft SQL Server 2008 به خاطر داشتن امنیت بالا، راحت بودن کار با آن، هزینه پایین دستیابی استفاده شده است. با استفاده از زبان برنامه‌نویسی Microsoft Visual Studio 2010 انجام شده است.

## ۸. برآورد زمان‌سفر با استفاده از خوشه‌بندی K-means

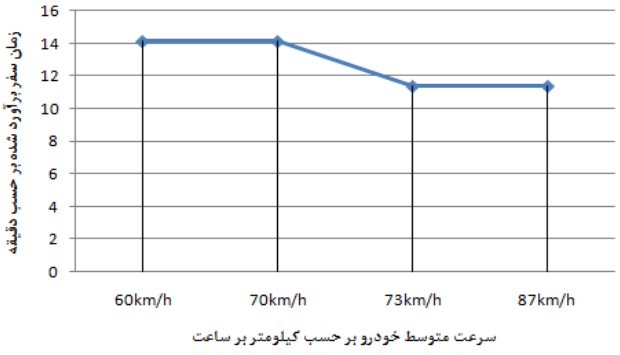
همان‌طور که گفته شده است در پیش‌پردازش داده‌ها تمام رکوردهای مشابه از نظر شرایط آب‌وهوایی، زمانی و تصادفات در خوشه‌های مجزایی قرار گرفته شده‌اند. سرعت متوسط هر خودرو تنها پارامتر قرار گرفته‌شده در هر خوشه می‌باشد که مشابه به یکدیگر نمی‌باشد. به‌علت آنکه داده‌های قرار گرفته‌شده در هر خوشه باید مشابه به یکدیگر باشند لذا در پردازش نهایی سرعت متوسط تمام خودروهای قرار گرفته شده در هر خوشه مشابه به یکدیگر می‌شوند. روند انجام پردازش‌نهایی در روش روش خوشه‌بندی K-means به این‌گونه می‌باشد که با مشخص نمودن تعداد دسته‌ها و یک حدآستانه به عنوان پارامترهای ورودی مورد نیاز، دسته‌بندی داده‌ها موجود در پایگاه داده انجام می‌گردد. نحوه دسته‌بندی این روش اینگونه است که با مشخص نمودن تعداد دسته‌ها، یکسری از اعداد (سرعت مربوط به خودروها) به‌صورت تصادفی درون هر دسته قرار می‌گیرند. پس از قرار گرفتن سرعت‌ها درون دسته‌ها، با استفاده از میانگین‌گیری، سرعت متوسط هر خوشه به‌دست می‌آید. در مرحله بعد با استفاده از حدآستانه ۳ معرفی شده به سیستم، تمامی سرعت‌ها که در همسایگی سرعت متوسط بوده و همین‌طور کمتر از حدآستانه می‌باشد، درون دسته‌ای که سرعت متوسط درون آن قرار داشته قرار خواهند گرفت و مجددا با استفاده از سرعت‌های قرار گرفته شده در هر دسته سرعت متوسط مربوط به آن دسته محاسبه می‌گردد. این کار تا جایی انجام می‌گیرد که تفاضل مقدار سرعت متوسط بدست آمده از هر دسته در هر مرحله نسبت به تکرار قبل از مقدار ۰/۰۰۱ کوچکتر باشد. معمولا این روش در تکرار اول و دوم به جواب بهینه می‌رسد.

زمان سفر برآورد شده توسط روش‌های K-means تحت شرایط ایده‌آل (بدون هیچ‌گونه تصادف‌جاده‌ای و شرایط‌آب‌وهوایی آفتابی در جدول ۴ آورده شده‌است. همان‌طور که در جدول ۴ مشاهده می‌شود در روش خوشه‌بندی K-means با توجه به روند تغییرات سرعت، تغییرات چندانی در زمان‌سفر حرکت خودرو به‌وجود نمی‌آید. به دلیل آنکه تعداد خوشه‌ها به عنوان ورودی در الگوریتم K-means دریافت می‌شود، لذا نسبت به ایجاد خوشه‌های مختلف، این الگوریتم حساس نمی‌باشد و نمی‌تواند به ازای افزایش و یا کاهش تغییرات سرعت خودرو، زمان‌سفر متناسب با سرعت‌های مختلف را برآورد نماید. میزان تاثیر بالاترین سرعت متوسط حرکت خودرو نسبت به پایین‌ترین سرعت متوسط حرکت خودرو در زمان‌سفر برآورد شده تحت

شرایط ایده‌آل (شرایط آب‌وهوایی آفتابی، بدون اوج و بدون تصادفات) ۳ دقیقه می‌باشد. در شکل ۳ روند تغییرات سرعت و زمان‌سفر با استفاده از الگوریتم K-means نشان داده شده است. روند تغییرات زمان‌سفر برآورد شده توسط روش خوشه‌بندی K-means در بازه‌های زمانی مختلف در جدول ۵ آورده شده‌است. همان‌طور که از جدول ۵ مشاهده می‌شود در گروه‌زمانی ۱و ۴ که زمان‌اوج حرکت خودرو می‌باشد، تردد خودروها به کندی انجام می‌گیرد و سرعت حرکت خودرو کاهش می‌یابد و لذا زمان رسیدن به شهر مقصد نسبت به شرایط بدون اوج بیشتر می‌گردد. روند تغییرات سرعت در زمان اوج و آب‌هوای بارندگی در مسیر قائمشهر به ساری در دامنه ۴۰ تا ۶۰ کیلومتر بر ساعت می‌باشد و در نتیجه تغییرات زمان‌سفر در بازه ۱۶/۱۸ تا ۲۱/۶۴ دقیقه می‌باشد. شکل ۴ بیانگر تغییرات زمان‌سفر برآورد شده با استفاده از روش خوشه‌بندی K-means در ساعت‌های مختلف روز می‌باشد.

جدول ۴. نمونه ای از زمان سفر برآورد شده در روش K-means تحت شرایط سرعت‌های مختلف

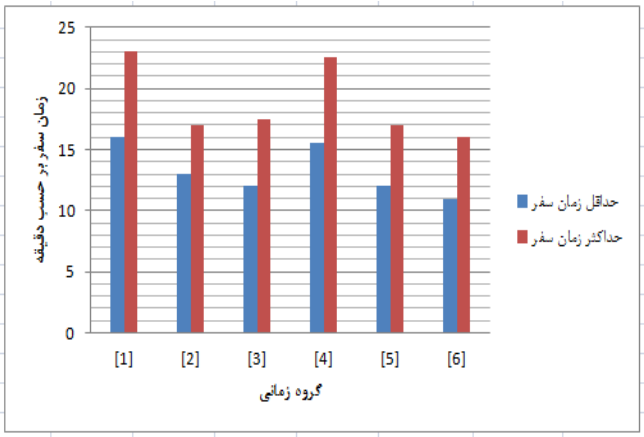
| زمان برآورد شده (بر حسب دقیقه) | سرعت متوسط خودرو در کل مسیر (بر حسب کیلومتر بر ساعت) | مقصد | مبدا |
|---|---|---|---|
| ۱۴/۱۳ | ۶۰ | قائمشهر | ساری |
| ۱۴/۱۳ | ۷۰ | قائمشهر | ساری |
| ۱۱/۳۶ | ۷۳ | قائمشهر | ساری |
| ۱۱/۳۶ | ۸۷ | قائمشهر | ساری |




شکل۳. زمان سفر برآورد شده با استفاده از روش خوشه‌بندی K-means در شرایط سرعت‌های مختلف

جدول ۵. بیانگر زمان‌سفر برآورد شده با استفاده از روش خوشه بندی K-means تحت شرایط آب‌وهوای بارندگی و زمان اوج

| زمان برآورد شده (دقیقه) | گروه‌زمانی | آب وهوا | سرعت متوسط خودرو (کیلومتر بر ساعت) | کد مسیر |
|---|---|---|---|---|
| ۲۱/۶۴ | ۱ | بارندگی | ۴۰ | ۱ |
| ۱۸/۶۴ | ۴ | بارندگی | ۴۸ | ۱ |
| ۱۶/۱۸ | ۱ | بارندگی | ۵۸ | ۱ |
| ۱۴/۱۳ | ۵ | آفتابی | ۶۰ | ۱ |
| ۱۴/۱۳ | ۳ | آفتابی | ۷۰ | ۱ |



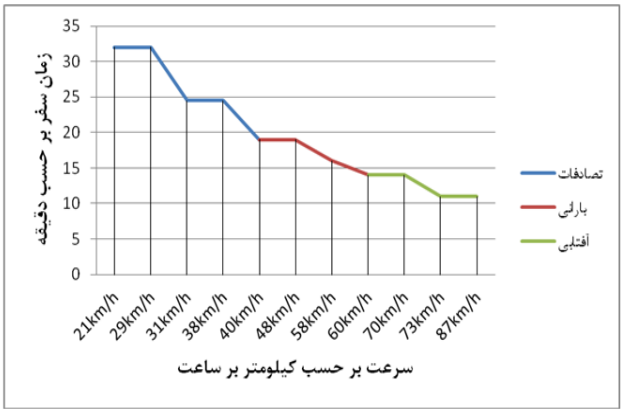

شکل ۵. زمان سفر برآورد شده با استفاده از روش خوشه‌بندی K-means در شرایط زمانی و آب‌وهوای مختلف

در شرایط تصادفات جاده‌ای و آب‌وهوای نامساعد نسبت میزان سرعت متوسط خودرو نسبت به شرایط ایده‌آل (شرایط آب وهوای آفتابی و بدون تصادف)، به‌علت توقف خودرو بشدت کاهش یافته و کاهش سرعت سبب طولانی‌تر شدن مدت زمان‌سفر می‌گردد و زمان سفر برآورد شده توسط الگوریتم K-means با سرعت متوسط ۲۵ کیلومتر بر ساعت، ۳۲/۴۸ دقیقه در مسیر ساری به قائمشهر برآورد شده است. می‌توان نتیجه گرفت که برآورد زمان‌سفر با استفاده از خوشه‌بندی K-means همانند خوشه‌بندی اتصال‌منفرد، اتصال‌کامل و روش خوشه‌بندی افزایشی، تاثیر شرایط تصادفات جاده‌ای در تاخیر زمان رسیدن به مقصد بیشتر از شرایط بارندگی و شرایط اوج‌زمانی می‌باشد. در روش خوشه‌بندی K-means چون فرآیند مقایسه بین یک‌به‌یک پارامتر سرعت متوسط خودروها انجام نمی‌گیرید و تعداد خوشه‌ها از قبل معلوم می‌باشد لذا زمان‌سفر برآورد شده توسط مدل خوشه‌بندی K-means نسبت به سرعت‌های مختلف دارای

اختلافات زمانی کم و یا صفر می‌باشد. شکل ۶ زمان سفر برآورد شده با استفاده از روش خوشه‌بندی K-means در شرایط تصادفات، آب‌وهوای بارندگی و ایده‌آل را نشان می‌دهد.

### ۹. تفسیر نتایج حاصل از الگوریتم K-means

با توجه به استفاده از روش خوشه‌بندی K-means، می‌توان تفسیرهای مختلفی را از نتایج حاصل از برآورد زمان‌سفر داشت. ذکر این نکته ضروری است که تفسیر نتایج که آخرین مرحله فرآیند داده کاوی است، بایستی توسط متخصص حوزه مورد مطالعه صورت پذیرد. با استفاده از روش‌های خوشه‌بندی K-means نشان داده شده است که در شرایط تصادفات‌جاده‌ای با بیشترین مقدار کاهش سرعت‌متوسط حرکت خودرو، با مقدار ۵۰درصد نسبت به شرایط ایده‌آل(آب‌وهوای آفتابی، بدون‌تصادف و بدون زمان‌اوج) در رسیدن به مقصد دارای تاخیرزمانی می‌باشد و شرایط آب‌وهوای بارندگی و زمان اوج با مقدار ۲۰درصد در رده‌های بعدی قرار دارند. روند تغییرات مدت‌زمان سفر را با توجه به تغییرات سرعت، تغییرات آب‌وهوایی و زمان‌های مختلف به‌دست آورد. جدول ۶ بیانگر میزان افزایش زمان‌سفر مسیر ساری به قائمشهر تحت شرایط تصادفات جاده‌ای، آب‌وهوای بارانی، زمان‌اوج و شرایط‌ایده‌آل را نشان می‌دهد. با توجه به زمان‌برآورد شده توسط روش‌های خوشه‌بندی بطور خلاصه می‌توان گفت که با توجه به بدیع بودن نتایج حاصل از روش‌های داده‌کاوی، می‌توان در کنار یک متخصص حوزه حمل‌ونقل و ترافیک‌شهری به تفسیرهای ارزشمندی در مورد خروجی‌های حاصل پرداخت.




شکل ۶. زمان سفر برآورد شده با استفاده از روش خوشه‌بندی K-means در شرایط تصادفات، آب‌وهوای بارندگی و ایده‌آل

جدول ۶. مقایسه تاثیر پارامتر های موثر زمان‌سفر

| درصدافزایش زمان سفر با توجه به شرایط ایده‌آل | بازه مدت زمان سفر (دقیقه) | بازه‌سرعت حرکت‌خودرو (کیلومتر بر ساعت) | عوامل برون‌جاده‌ای |
|---|---|---|---|
| ۰ | [10 15] | [60 90] | شرایط ایده‌آل |
| حداکثر ٪۲۰ | [15 20] | [40 60] | زمان‌اوج |
| حداکثر ٪۲۰ | [15 20] | [40 60] | آب‌وهوایی مختلف |
| حداکثر ٪۵۰ | [25 40] | [20 40] | تصادفات‌جاده‌ای |

### ۹. شبکه عصبی مصنوعی

طراحی شبکه‌عصبی یکی از بخش های مهم و تاثیر گذار بر عملکرد و نتایج شبکه عصبی می‌باشد. معماری شبکه عصبی به عواملی از جمله تعداد لایه‌های پنهان، نرون‌های هر لایه، توابع انتقال، و توابع آموزش وابسته می‌باشند. جدول ۷ بیانگر پارامترهای تاثیرگذار و سطوح مختلف هر یک از این پارامترها را در این تحقیق نشان می‌دهد. لازم به ذکر است است. هر چقدر تعداد لایه‌های پنهان و تعداد نرون‌ها بیشتر گردد به‌علت فرآیند آموزش بیش از اندازه شبکه خطای شبکه‌عصبی در پیش‌بینی بیشتر می‌گردد لذا تعداد لایه‌های پنهان در این تحقیق یک در نظر گرفته می‌شود. توابع انتقال بر دو نوع خطی و غیرخطی می‌باشند. در این تحقیق هدف برآورد زمان‌سفر نسبت به یکسری از پارامترهای ورودی می‌باشد و ممکن است جواب برآورد شده الزاما در میان جواب‌های معلوم نباشد. شبکه‌های عصبی در این موارد، باید به‌گونه‌ای باشند که قابلیت تولید مجموعه نامتناهی از پاسخ‌ها را داشته باشند. لذا از توابع خطی به‌علت محدود کردن پارامترهای ورودی استفاده نمی‌شود و از توابع غیرخطی که سه نوع توابع Tansig ، Radbas و Logsig می‌باشند استفاده می‌شود و با آزمون سعی و خطا، مناسب ترین تابع انتقال انتخاب می‌شود. از توابع آموزش Trainlm وTraingdx به‌علت آنکه در تعداد تکرارهای کمتری نسبت به سایر روش گرادیان شیب و نیوتن همگرا می‌شود. Traingdx وTrainlm سریعترین الگوریتم آموزش شبکه‌های عصبی با تعداد پارامترهای موثر متوسط می‌باشند. سرعت یادگیری و سرعت حرکت در هر تکرار تغییرپذیر است تا آنکه درصد زمان یادگیری را کاهش دهد.

برای دستیابی به معماری مناسب شبکه‌عصبی، تعدادی از عوامل تاثیرگذار جدول ۶ با هم ترکیب می‌شود تا آنکه با آزمون سعی‌وخطا بتوان به بهترین معماری شبکه‌عصبی مصنوعی برای پیش‌بینی و برآورد زمان‌سفر دست پیدا نمود. برای این منظور با

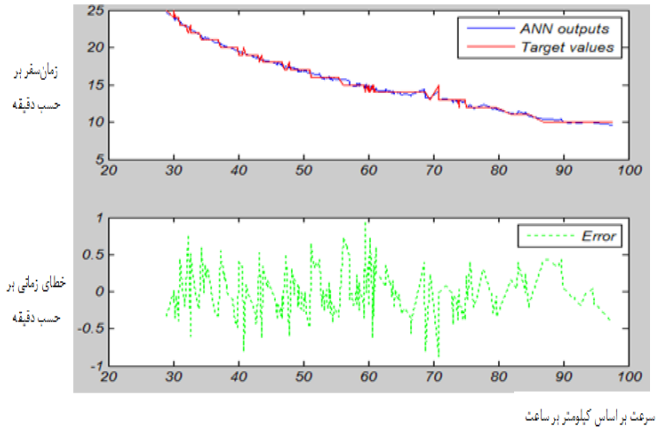
توجه به جدول ۶ و با در نظر گرفتن مقدار سطح هر پارامتر حدود ۳۰=۲×۳×۵×۱ روش برای فهمیدن بهترین معماری شبکه‌عصبی وجود دارد که با تست بر روی آنها پی به معماری درست برآورد زمان‌سفر با استفاده از شبکه‌عصبی برده شده است. معماری اصلی شبکه‌عصبی با یک لایه پنهان و تعداد ۸ نرون در هر لایه در نظر گرفته می‌شود و تابع انتقال tansig و تابع آموزش trainlm می‌باشد. با مشاهده نمودار خطای زمانی شکل ۷، مشاهده می‌گردد که خطای برآورد زمان‌سفر شبکه عصبی با مقدار واقعی بین اعداد [1 -1] می‌باشد. مقدار تابع همبستگی داده‌های آموزش برابر با ۰/۹۹۷ و داده‌های اعتبارسنجی نزدیک به ۰/۹۹۵ و داده‌های تست نزدیک به ۰/۹۹۳ می‌باشد و طبق تابع مجموع مربعات میانگین خطای نسبی در تکرار ۱۷ به جواب بهینه می‌رسد. شکل ۸ بیانگر تعداد تکرارهایی است که شبکه عصبی با معماری فوق به جواب بهینه می‌رسد.

زمان سفر برآورد شده توسط شبکه‌عصبی مصنوعی تحت شرایط ایده‌آل (بدون هیچ‌گونه تصادف‌جاده‌ای و شرایط‌آب‌وهوایی آفتابی در شکل ۹ آورده شده‌است. همان‌طور که در شکل ۹ مشاهده می‌شود در مسیر قائمشهر به ساری، هر چقدر میزان سرعت حرکت خودرو بیشتر باشد زمان رسیدن به شهر مقصد کمتر می‌شود.

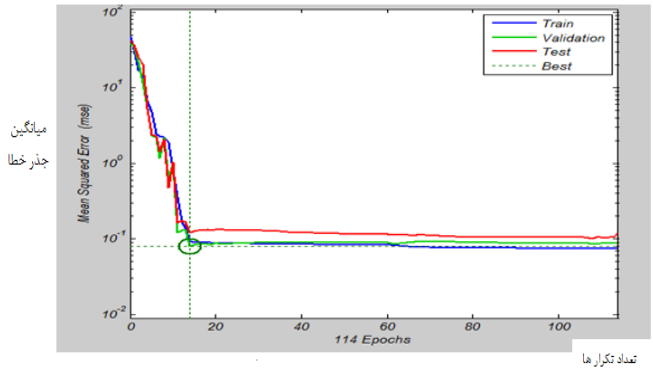
شکل ۱۰ بیانگر روند تغییرات زمان‌سفر تحت شرایط آب‌وهوایی مختلف در مسیر قائمشهر به ساری می‌باشد. افزایش میزان شدت بارندگی و سرعت باد باعث کند شدن حرکت خودرو و افزایش زمان‌سفر در مسیر می‌شود. همانطور که شکل ۱۰ نشان می‌دهد، با افزایش میزان شدت بارش و سرعت یافتن باد، زمان‌سفر یک روند صعودی خواهد داشت و زمان‌سفر افزایش می‌یابد.

جدول۷. اجزاء و سطوح هر جزء بکار گرفته شده در شبکه عصبی

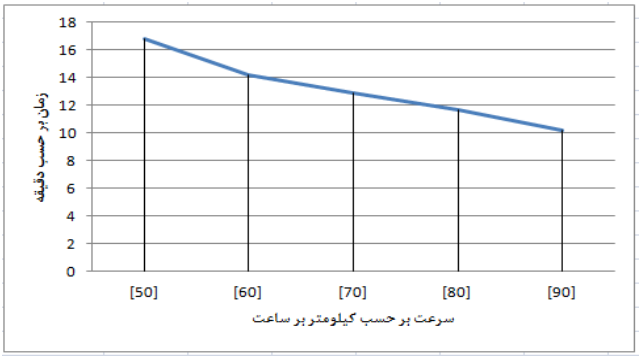
| پارامترهای طراحی | سطوح مختلف تاثیرگذاری | تعداد سطوح |
|---|---|---|
| لایه های پنهان | ۱ لایه | ۱ |
| نرون های هر لایه | متغیر بین ۷ تا ۱۲ | ۵ |
| تابع تبدیل | Tansig,Radbas,Logsig | ۳ |
| تابع آموزش | Trainlm,Traingdx | ۲ |



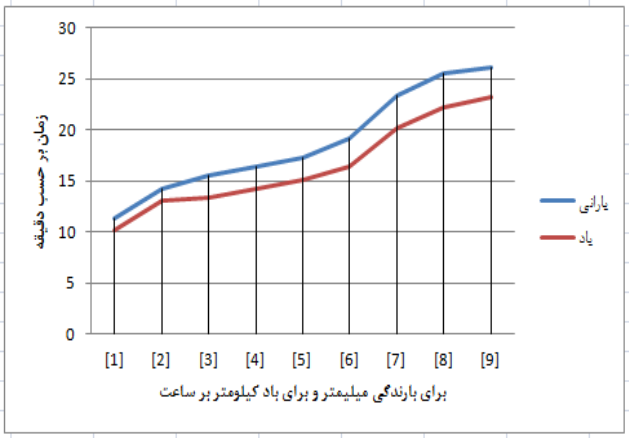

شکل۷. بیانگر بازه خطای شبکه عصبی بهینه‌ترین معماری




شکل۸. بیانگر تعدادتکرارهایی که شبکه عصبی به جواب بهینه می‌رسد




شکل ۹. روند تغیرات زمان‌سفر برآورد شده با استفاده از شبکه عصبی در سرعت‌های مختلف




شکل ۱۰. روند تغیرات زمان‌سفر برآورد شده با استفاده از شبکه عصبی در شرایط آب‌وهوای نامساعد

در شرایط تصادفات جاده‌ای و آب‌وهوای بارانی نسبت میزان سرعت متوسط خودرو نسبت به شرایط ایده‌آل (شرایط آب وهوایی آفتابی و بدون تصادف)، به‌علت متوقف شدن حرکت خودرو کاهش می‌یابد. کاهش حرکت خودرو سبب بالا رفتن زمان‌سفر حرکت خودرو می‌شود. زمان سفر برآورد شده توسط روش‌شبکه عصبی با شرایط تصادف جاده‌ای و آب‌وهوای بارانی و سرعت متوسط ۲۵ کیلومتر بر ساعت، ۲۸/۲۳ دقیقه در مسیر ساری به قائمشهر برآورد شده است.

## ۱۰. ارزیابی و مقایسه خوشه‌بندی K-means و شبکه‌عصبی با داده‌های واقعی

به‌منظور برآورد دقت حاصل از روش خوشه‌بندی K-means از یکسری داده‌های واقعی از مجموعه داده‌های موجود در پایگاه داده استفاده خواهیم نمود. داده‌های انتخاب شده به صورت تصادفی از یکسری داده‌های واقعی خودروهایی که مسیر بین شهرهای ساری به قائمشهر و بابل به قائمشهر را طی نموده‌اند، انتخاب شده‌اند. برای محاسبه دقت نتایج بدست آمده از روش‌های میانگین مربع خطا (MSE) و میانگین ریشه مربع خطا (RMSE) استفاده شده است رابطه (۲) و (۳).

$$MSE = \frac{1}{n}\left[\sum_{i=1}^{n}\left(X_i - X\right)^2\right] \qquad (۲)$$

$$RMSE = \sqrt{\frac{1}{n}\left[\sum_{i=1}^{n}\left(X_i - X\right)^2\right]} \qquad (۳)$$

زمان سفر برآورد شده با استفاده از روش شبکه‌عصبی مصنوعی و خوشه‌بندی K-means در بازه‌زمانی مختلف و بدون اوج و تحت شرایط سرعت‌های مختلف آب‌وهوای آفتابی در مسیر ساری به قائمشهر با استفاده از داده‌های واقعی مقایسه شدند و روش شبکه عصبی با استفاده از روش میانگین ریشه مربع خطا با دقت ۰/۰۰۲ با بالاترین مقدار دقت نسبت به داده‌های واقعی، زمان‌سفر را برآورد کرده است و روش‌های K-means با دقت ۰/۰۰۹ به‌ترتیب در رده بعدی قرار دارد. زمان سفر برآورد شده با استفاده از روش شبکه‌عصبی مصنوعی و خوشه‌بندی K-means، علاوه بر روش میانگین ریشه مربع خطا با استفاده از روش میانگین مربع خطا نیز ارزیابی شده‌اند. روش شبکه عصبی با استفاده از روش میانگین مربع خطا با دقت ۰/۰۴ با بالاترین مقدار دقت نسبت به داده‌های واقعی زمان‌سفر را برآورد کرده است و روش K-means با دقت ۰/۱۶ در بعد آن قرار دارد.

زمان‌سفر برآورد شده با استفاده از روش شبکه‌عصبی مصنوعی و خوشه‌بندی K-means در بازه‌زمانی اوج و تحت شرایط آب‌هوای بارندگی و زمان‌اوج در مسیر بابل به قائمشهر با استفاده از داده‌های واقعی مقایسه شدند. روش شبکه عصبی با استفاده از روش میانگین ریشه مربع خطا با دقت ۰/۰۰۱ با بالاترین مقدار دقت نسبت به داده‌های واقعی، زمان‌سفر را برآورد کرده است و روش خوشه‌بندی K-means با دقت‌های ۰/۰۳۲ در رده بعدی قرار دارد. زمان سفر برآورد شده با استفاده از روش شبکه‌عصبی مصنوعی و خوشه‌بندی K-means ، علاوه بر روش میانگین ریشه مربع خطا با استفاده از روش میانگین مربع خطا نیز ارزیابی شده‌اند. روش شبکه عصبی با استفاده از روش میانگین مربع خطا با دقت ۰/۰۹ با بالاترین مقدار دقت نسبت به داده‌های واقعی زمان‌سفر را برآورد کرده است و روش خوشه‌بندی K-means با ۱/۳۴ بعد آن قرار دارد.

در روش‌شبکه‌عصبی مصنوعی چون‌که داده‌های سرعت‌خودرو، بازه زمانی حرکت خودرو، شرایط آب‌وهوایی مسیر حرکت‌خودرو، شرایط تصادفات جاده‌ای و مسیر حرکت خودرو به‌صورت همزمان به درون شبکه وارد می‌شوند و شبکه عصبی مصنوعی طوری است که خطای شبکه نسبت به کل‌داده‌ها محاسبه می‌شود، پس از محاسبه خطای شبکه، وزن‌های هر پارامتر نسبت به خطای شبکه تغییر می‌یابند تا به جواب ایده‌آل برسند. فرآیند تکرار تغییر وزن‌های هر پارامتر در جهت رسیدن به جواب ایده‌آل سبب شده است که دقت روش شبکه‌عصبی مصنوعی نسبت به روش‌های خوشه‌بندی بیشتر گردد. روش‌های دیگر خوشه‌بندی

در برآورد زمان‌سفر همانند روش‌های خوشه‌بندی اتصال‌کامل، اتصال‌منفرد و افزایشی دارای دقت مشابه هستند. علت این امر در این است که در روش خوشه‌بندی اتصال‌کامل، اتصال‌منفرد و افزایشی با مقایسه مستقیم بین پارامترها فرآیند تشکیل خوشه‌ها حاصل می‌شود و در صورتیکه در روش خوشه‌بندی K-means مقایسه مستقیم بین پارامترها انجام نمی‌گیرد و تعداد خوشه‌ها با توجه به مقایسه بین پارامترها حاصل نمی‌شود، دقت روش K-means در میان روش‌های دیگر خوشه‌بندی در رتبه آخر می‌باشد.

## ۱۱. نتیجه‌گیری

برآورد زمان‌سفر در فرآیندهای برنامه‌ریزی سیستم‌های حمل‌ونقل، مورد توجه متخصصان حوزه‌های مختلف بوده که هر یک با بهره‌گیری از داده‌ها و ابزارهای تحلیلی خاص، سعی در مطالعه و برآورد زمان سفر داشته‌اند. برآورد زمان‌سفر به پارامترهای مختلف طبیعی و انسانی وابسته است. در این زمینه مطالعات بسیاری صورت گرفته که در اکثر آنها پارامترهای محدودی را در نظر گرفته‌اند و یا برای برآورد زمان‌سفر از مدل‌های تجربی استفاده شده است. هدف از این تحقیق برآورد زمان‌سفر تحت تاثیر شرایط مختلف آب‌وهوایی، تصادفات جاده‌ای و بازه‌های زمانی مختلف روز در مسیرهای استان مازندران بوده است، که تا بحال این کار صورت نگرفته است. در این تحقیق در فاز اول به بیان پارامترهای مختلف موثر در برآورد زمان‌سفر در مسیرهای استان مازندران پرداخته شد. با توجه به این که تعداد این پارامترها زیاد بوده و هر یک از این پارامترها تاثیرات متفاوتی در زمان‌سفر دارند، لذا بایستی با توجه به امکانات و شرایط‌جاده‌ای پارامترها استخراج شوند تا به منظور برآورد زمان سفر در روش‌های شبکه‌عصبی مصنوعی خوشه‌بندی وK-means استفاده شوند. داده‌های مورد استفاده در این تحقیق از متوسط سرعت‌حرکت خودرو، بازه‌زمانی حرکت خودرو، مدت زمان‌حرکت خودرو، شرایط آب‌وهوایی مسیر، شرایط تصادفات در مسیر حرکت خودرو در بازه‌زمانی سال‌های ۱۳۸۸ و ۱۳۹۰ استفاده شده است.

نتایج زیر از ارزیابی مدل های مورد استفاده به‌دست آمد.

در روش‌شبکه‌عصبی مصنوعی چون‌که داده‌های سرعت‌خودرو، بازه زمانی حرکت خودرو، شرایط آب‌وهوایی مسیر حرکت‌خودرو، شرایط تصادفات جاده‌ای و مسیر حرکت خودرو به‌صورت همزمان به درون شبکه وارد می‌شوند و شبکه عصبی مصنوعی طوری است که خطای شبکه نسبت به کل‌داده‌ها محاسبه می‌شود. پس از محاسبه خطای شبکه، وزن‌های هر پارامتر نسبت به خطای شبکه تغییر می‌یابند تا به جواب ایده‌آل برسند. فرآیند تکرار تغییر وزن‌های هر پارامتر در جهت رسیدن به‌جواب ایده‌آل سبب شده است که دقت روش شبکه‌عصبی مصنوعی نسبت به روش‌های خوشه‌بندی بیشتر گردد.

در میان چهار روش خوشه‌بندی اتصال‌کامل، اتصال‌منفرد، افزایشی و K-means دقت حاصل از سه روش خوشه‌بندی اتصال‌کامل، اتصال‌منفرد و افزایشی مشابه هست و علت این امر در این است که در روش خوشه‌بندی اتصال‌کامل، اتصال‌منفرد و افزایشی با مقایسه مستقیم بین پارامترها فرآیند تشکیل خوشه‌ها حاصل می‌شود و در صورتی‌که در روش خوشه‌بندی K-means مقایسه مستقیم بین پارامترها انجام نمی‌گیرد و تعداد خوشه‌ها با توجه به مقایسه بین پارامترها حاصل نمی‌شود.

بررسی زمان‌سفر با استفاده از پارامترهای برون جاده‌ای در مسیر قائمشهر به بابل و ساری به قائمشهر نشان می‌دهد شرایط تصادفات جاده‌ای بیشترین تاخیر را بعلت کاهش زیاد سرعت متوسط حرکت خودرو سبب می‌شوند و شرایط آب‌وهوایی و زمان اوج حرکت خودرو در رتبه‌های بعدی قرار دارند.

## مراجع


[۱] موسوی ع، ارائه یک روش مبتنی بر داده کاوی مکانی به منظور تحلیل خطرپذیری خشکسالی ،پایان نامه کارشناسی ارشد، دانشکده فنی، دانشگاه تهران، تهران، ایران، ۱۳۸۸.

[۲] زهرایی ب.، گزارش نهایی تدوین مدل خوشه بندی اطلاعات و سیگنال های هواشناسی با هدف پیش بینی دوره های کم بارش، پروژه تحقیقاتی، سازمان مجری دانشگاه تهران، ۱۳۸۸.

[۳] کانتاردزیک م.، داده کاوی، ترجمه: امیر علیخانزاده، نشر علوم رایانه، بابل، ایران، ۱۳۸۵.

[۴] ناصری م. ا.، برادران و.، بررسی عوامل موثر بر زمان سفر در سیستم حمل و نقل عمومی و پیش بینی زمان سفر مورد کاوی: سیستم اتوبوسرانی شهر تهران، پژوهشنامه حمل و نقل، شماره ۳، صفحات ۲۳۲-۲۱۹، ۱۳۸۷.

[5] Wen-ting, L., Zhi-jian, W., Qin, Y., Travel time prediction of multi-source historical data fusion, Journal of Advanced Materials and Computer Science, V. ۴۷۴-۴۷۶, PP. ۷۷۷-۷۸۱, ۲۰۱۱.

[6] Deb Nath, R., Lee, H., Chowdhury, N., Chang, J., Modified K-Means Clustering for Travel Time Prediction Based on Historical Traffic Data, Proceedings of the 14th international conference on Knowledge-based and intelligent information and engineering systems, PP. ۵۲۱-۵۱۱, ۲۰۱۰.

[7] Chowdhury, Ch., Deb Nath, R., Lee, H., Chang, J., Development of an Effactive Travel Time Predication Method Using Modified Moving Average Approach, Proceedings of the 13th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, PP. ۱۳۰-۱۳۸, ۲۰۰۹.

[8] Chowdhury, N., Carson, K., Leung, S., Improved Travel Time Prediction Algorithms for Intelligent Transportation Systems, Proceedings of the 15th international conference on Knowledge-based and intelligent information and engineering systems, V. ۲, PP. ۳۵۵-۳۶۵, ۲۰۱۱.

[9] K. Jain and R. C. Dubes, *Algorithms for Clustering Data*, Prentice Hall, Englewood Cliffs, 1988.

[10] P. H. A. Sneath and R. R. Sokal, *Numerical Taxonomy*, Freeman, San Francisco, 1973.

[11] B. King, *Step-Wise Clustering Procedures*, Journal of the American Statistical Association,69:86-101, 1967.

[12] Murtagh, A survey of Recent Advances in Hierarchical Clustering Algorithms Which Use Cluster centers, The Computer Journal, 26:354-359, 1984.

[13] http://fa.wikipedia.org/wiki/

[14] Kantardzic, M., *Data Mining: Concepts, Models, Methods, and Algorithms*, copy right 2011 by Institute of Electrical and Electronics Engineers, Published simulataneously in Canada.

[15] http://forum.persiantools.com/t71559.html

[16] Sumit, G., Gyanendra, K., A Novel Method for Shelf Life Detection of Processed Cheese Using Cascade Single and Multi Layer Artificial Neural Network Computing Models, ARPN Journal of Systems and Software, Vol. 2, No. 2, PP. 51-83, 2012.

[17] Shamseldin, A., Ahmed, E., Nasr, K., O'Connor, M., Comparison of different forms of the Multi-layer Feed-Forward Neural Network method used for river flow_forecasting, Hydrology and Earth System Sciences, Vol. ۶, No. ۴, PP. ۶۷۱-۶۸۴, ۲۰۰۲.