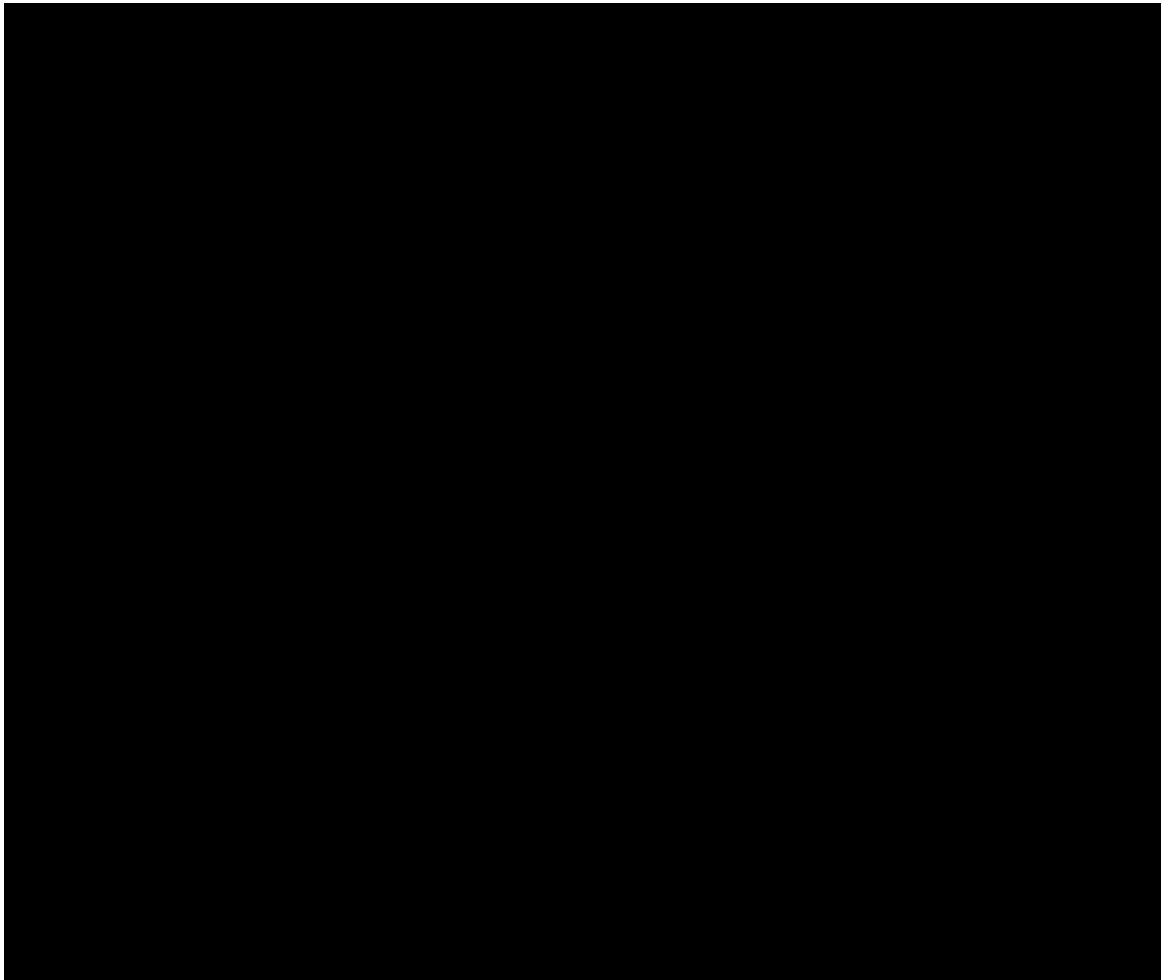






دانشکده مهندسی عمران

پایان نامه کارشناسی ارشد مهندسی عمران – راه و ترابری

پیش بینی جریان ترافیک با استفاده از سری های زمانی و یادگیری عمیق

نگارنده: مجتبی محمدزاده

استاد راهنما

دکتر عبدالاحد چوپانی

بهمن ۱۴۰۰

در این صفحه صورت جلسه دفاع را قرار دهید. لازم است پس از صحافی این صفحه مجدداً توسط دانشکده مهر گردد و استاد راهنما با امضای خود اصلاحات پایان نامه را تایید کند.

تقدیم اثر

پروردگار!:

نه می توانم موهایشان را که در راه عزت من سفید شد، سیاه کنم و نه برای دستهای پینه بسته شان که ثمره تلاش برای افتخار من است، مرهمی دارم. پس توفیقم ده که هر خطه شکر گزارشان باشم و ثانیه های عمرم را در عصای دست بودنشان بگذرانم.

شکر و سپاس خدا را که بزرگترین امید و یاور در خطه خطه زندگیم است.

شکر و قدردانی

از استاد کرامیم جناب آقای دکتر عبدالاحد چوپانی بسیار سپاسگذارم چرا که بدون راهنمایی های ایشان تاین این پایان نامه
بسیار مشکل می نمود.

از دوستان و بهکلاسی های کرامی که ایام خوشی را در کنار هم سپری کردیم، شکر می کنم و از دوست خوبم مهندس داوود
فاضلی که در کد نویسی متلب در این پایان نامه من را یاری نمودند شکر می کنم.

تهدنام

اینجانب **مجتبی محمدزاده** دانشجوی دوره کارشناسی ارشد رشته راه و ترابری دانشکده عمران دانشگاه صنعتی شاهرود نویسنده پایان نامه پیش‌بینی جریان ترافیک با استفاده از سری‌های زمانی و یادگیری عمیق تحت راهنمایی آقای دکتر چوپانی متعهد می‌شوم.

- تحقیقات در این پایان‌نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است.
- مطالب مندرج در پایان‌نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می‌باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان‌نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان‌نامه رعایت می‌گردد.
- در کلیه مراحل انجام این پایان‌نامه، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این پایان‌نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری، ضوابط و اصول اخلاق انسانی رعایت شده است.

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه ای، نرم افزار ها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می‌باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود. استفاده از اطلاعات و نتایج موجود در پایان‌نامه بدون ذکر مرجع مجاز نمی‌باشد.

چکیده

پیش‌بینی جریان ترافیک در چند دهه گذشته توجه زیادی را به خود جلب کرده است. این موضوع هنوز یک موضوع چالش برانگیز برای محققان ترافیک و حمل و نقل می‌باشد. باتوجه به ویژگی‌های تصادفی جریان ترافیک پیش‌بینی دقیق ترافیک کار ساده‌ای نخواهد بود. از این جهت برای حل این موضوع تکنیک‌های زیادی برای مدلسازی و پیش‌بینی جریان ترافیک به کار گرفته شده‌است. در پژوهش حاضر از روش پارامتریک، ARIMA (AutoRegressive Integrated Moving Average)، روش غیرپارامتریک، شبکه عصبی مصنوعی (Artificial neural network) و روش یادگیری عمیق، LSTM (Long short-term memory) برای پیش‌بینی حجم تردد ترافیک جاده‌ای برای ۶ مسیر رفت و برگشت منتهی به شهر مشهد استفاده گردیده‌است و نتایج این سه روش با یکدیگر مقایسه شده‌اند. نتایج، نشان دهنده عملکرد بهتر روش یادگیری عمیق LSTM نسبت به دو روش دیگر می‌باشد. مقادیر شاخص‌های RMSE، MAPE و AADT برای این سه روش اختلاف قابل توجهی نسبت به یکدیگر دارند. اختلاف مقادیر شاخص‌های RMSE، MAPE و AADT مدل ARIMA با مدل یادگیری عمیق LSTM به ترتیب ۰٫۷۲، ۲۴٪ و ۲۳٪ می‌باشد. همچنین اختلاف مقادیر این شاخص‌ها برای مدل شبکه عصبی مصنوعی و یادگیری عمیق LSTM به ترتیب ۰٫۱۸، ۱۱٪ و ۶٪ می‌باشد که توانایی بالا روش‌های یادگیری عمیق را نسبت به روش‌های پارامتریک و غیرپارامتریک نشان می‌دهد.

کلمات کلیدی: پیش‌بینی حجم تردد، سری‌های زمانی، ARIMA، LSTM، شبکه عصبی مصنوعی، یادگیری عمیق

لیست مقالات مستخرج از پایان نامه

۱- پیش بینی حجم مسافران ایستگاه راه آهن با استفاده از سری زمانی فصلی **SARIMA** : مطالعه موردی مشهد

۲-

۳-

فهرست مطالب

۵	فهرست جداول
۹	فهرست اشکال
۱	فصل ۱: مقدمه
۲	۱ - ۱ مقدمه
۲	۱ - ۲ تعریف مساله
۳	۱ - ۳ اهمیت پیش بینی و حوزه کاربرد
۴	۱ - ۴ اهداف تحقیق
۵	۱ - ۵ فرضیه ها
۶	۱ - ۶ روش و مراحل تحقیق
۶	۱ - ۶ - ۱ روش ARIMA
۶	۱ - ۶ - ۲ روش شبکه عصبی و LSTM
۷	۱ - ۷ ضرورت حل مسئله
۸	۱ - ۸ چالش های تحقیق
۸	۱ - ۹ ساختار پژوهش
۹	فصل ۲: مبانی نظری و پیشینه تحقیق
۱۰	۲ - ۱ مقدمه
۱۰	۲ - ۲ مبانی نظری
۱۰	۲ - ۲ - ۱ سری های زمانی
۱۱	۲ - ۲ - ۲ روش ARIMA
۱۳	۲ - ۲ - ۳ مدل شبکه عصبی مصنوعی و مدل LSTM

۲ - ۳ پیشینه تحقیق..... ۲۱

۲ - ۳ - ۱ روش پارامتریک (ARIMA)..... ۲۱

۲ - ۳ - ۲ مقایسه روش‌های پارامتریک و غیرپارامتریک..... ۲۲

۲ - ۳ - ۳ یادگیری عمیق..... ۲۴

۲ - ۳ - ۴ جمع بندی..... ۲۷

فصل ۳: روش تحقیق ۲۹

۳ - ۱ مقدمه..... ۳۰

۳ - ۲ روش ARIMA و نرم افزار SPSS..... ۳۰

۳ - ۲ - ۱ مراحل انجام پیش بینی در نرم افزار SPSS به روش ARIMA..... ۳۱

۳ - ۳ روش شبکه عصبی و یادگیری عمیق LSTM..... ۳۲

۳ - ۳ - ۱ برنامه‌نویسی شبکه عصبی مصنوعی و یادگیری عمیق LSTM در نرم‌افزار

MATLAB..... ۳۳

۳ - ۴ نتیجه‌گیری..... ۳۵

فصل ۴: مدل‌سازی و تحلیل نتایج ۳۷

۴ - ۱ مقدمه..... ۳۸

۴ - ۲ معرفی مسیر های مورد ارزیابی..... ۳۸

۴ - ۳ گردآوری داده‌ها..... ۴۲

۴ - ۴ مدل‌سازی و پیش‌بینی به روش ARIMA..... ۴۳

۴ - ۵ مدل‌سازی و پیش‌بینی شبکه عصبی مصنوعی..... ۶۳

۴ - ۶ مدل‌سازی و پیش‌بینی یادگیری عمیق LSTM..... ۸۲

۴ - ۷ اعتبارسنجی..... ۱۰۱

۴ - ۷ - ۱ مقایسه مدل‌سازی‌ها با استفاده از خطاهای MSE و RMSE..... ۱۰۲

۲ - ۷ - ۴ مقایسه مدلسازی‌ها با استفاده از خطا MAPE ۱۰۴

۳ - ۷ - ۴ مقایسه مدلسازی‌ها با استفاده از AADT ۱۰۶

فصل ۵: نتیجه‌گیری و پیشنهادات ۱۱۵

۵ - ۱ مقدمه ۱۱۶

۵ - ۲ نتایج پیش‌بینی ۱۱۶

۵ - ۳ مقایسه سه مدل پیش‌بینی ۱۱۷

۵ - ۴ پیشنهادات ۱۱۸

مراجع ۱۲۰

پیوست ۱۲۳

فهرست جداول

جدول ۱-۴ معرفی و بازه زمانی برداشت داده‌های مسیرهای استان خراسان رضوی.....	۴۲
جدول ۲-۴ اطلاعات کلی از داده‌های مورد ارزیابی و پیش‌بینی.....	۴۲
جدول ۳-۴ پارامترهای مدل برای محور مشهد - سرخس با حذف ضریب ثابت.....	۴۴
جدول ۴-۴ مقادیر خطا در محور مشهد- سرخس.....	۴۵
جدول ۵-۴ پارامترهای مدل برای محور سرخس - مشهد با حذف ضریب ثابت.....	۴۶
جدول ۶-۴ مقادیر خطا در محور سرخس - مشهد.....	۴۷
جدول ۷-۴ پارامترهای مدل برای مسیر مشهد - چناران با حذف ضریب ثابت.....	۴۷
جدول ۸-۴ مقادیر خطا در محور مشهد- چناران.....	۴۸
جدول ۹-۴ پارامترهای مدل برای محور چناران - مشهد با حذف ضریب ثابت.....	۴۹
جدول ۱۰-۴ مقادیر خطا در محور چناران - مشهد.....	۵۰
جدول ۱۱-۴ پارامترهای مدل برای مسیر مشهد - فریمان با حذف ضریب ثابت.....	۵۰
جدول ۱۲-۴ مقادیر خطا در محور مشهد- فریمان.....	۵۱
جدول ۱۳-۴ پارامترهای مدل برای محور فریمان - مشهد با حذف ضریب ثابت.....	۵۲
جدول ۱۴-۴ مقادیر خطا در محور فریمان - مشهد.....	۵۳
جدول ۱۵-۴ پارامترهای مدل برای مسیر مشهد - ملک‌آباد با حذف ضریب ثابت.....	۵۳
جدول ۱۶-۴ مقادیر خطا در محور مشهد- ملک‌آباد.....	۵۴
جدول ۱۷-۴ پارامترهای مدل برای محور ملک‌آباد - مشهد با حذف ضریب ثابت.....	۵۵
جدول ۱۸-۴ مقادیر خطا در محور ملک‌آباد - مشهد.....	۵۶
جدول ۱۹-۴ پارامترهای مدل برای مسیر مشهد - سدکارده با حذف ضریب ثابت.....	۵۷
جدول ۲۰-۴ مقادیر خطا در محور مشهد- سدکارده.....	۵۸
جدول ۲۱-۴ پارامترهای مدل برای محور سدکارده - مشهد با حذف ضریب ثابت.....	۵۸
جدول ۲۲-۴ مقادیر خطا در محور سدکارده - مشهد.....	۵۹
جدول ۲۳-۴ پارامترهای مدل برای مسیر مشهد - کلات با حذف ضریب ثابت.....	۶۰
جدول ۲۴-۴ مقادیر خطا در محور مشهد- کلات.....	۶۱
جدول ۲۵-۴ پارامترهای مدل برای محور کلات - مشهد با حذف ضریب ثابت.....	۶۱
جدول ۲۶-۴ مقادیر خطا در محور کلات - مشهد.....	۶۲
جدول ۲۷-۴ زمان مدل‌سازی در مدل شبکه عصبی مصنوعی.....	۶۳
جدول ۲۸-۴ مقادیر خطا در محور مشهد - سرخس.....	۶۵
جدول ۲۹-۴ مقادیر خطا در محور سرخس - مشهد.....	۶۷
جدول ۳۰-۴ مقادیر خطا در محور مشهد - چناران.....	۶۸

جدول ۳۱-۴	مقادیر خطا در محور چناران - مشهد	۷۰
جدول ۳۲-۴	مقادیر خطا در محور مشهد - فریمان	۷۱
جدول ۳۳-۴	مقادیر خطا در محور فریمان - مشهد	۷۳
جدول ۳۴-۴	مقادیر خطا در محور مشهد - ملک آباد	۷۴
جدول ۳۵-۴	مقادیر خطا در محور ملک آباد - مشهد	۷۶
جدول ۳۶-۴	مقادیر خطا در محور مشهد - سدکارده	۷۷
جدول ۳۷-۴	مقادیر خطا در محور سدکارده - مشهد	۷۹
جدول ۳۸-۴	مقادیر خطا در محور مشهد - کلات	۸۰
جدول ۳۹-۴	مقادیر خطا در محور کلات - مشهد	۸۲
جدول ۴۰-۴	زمان مدلسازی در شبکه LSTM	۸۳
جدول ۴۱-۴	مقادیر خطا در محور مشهد - سرخس	۸۵
جدول ۴۲-۴	مقادیر خطا در محور سرخس - مشهد	۸۶
جدول ۴۳-۴	مقادیر خطا در محور مشهد - چناران	۸۸
جدول ۴۴-۴	مقادیر خطا در محور چناران - مشهد	۸۹
جدول ۴۵-۴	مقادیر خطا در محور مشهد - فریمان	۹۱
جدول ۴۶-۴	مقادیر خطا در محور فریمان - مشهد	۹۲
جدول ۴۷-۴	مقادیر خطا در محور مشهد - ملک آباد	۹۴
جدول ۴۸-۴	مقادیر خطا در محور ملک آباد - مشهد	۹۵
جدول ۴۹-۴	مقادیر خطا در محور مشهد - سدکارده	۹۷
جدول ۵۰-۴	مقادیر خطا در محور سدکارده - مشهد	۹۸
جدول ۵۱-۴	مقادیر خطا در محور مشهد - کلات	۱۰۰
جدول ۵۲-۴	مقادیر خطا در محور کلات - مشهد	۱۰۱
جدول ۵۳-۴	خطای MSE و RMSE برای مدلسازی روزانه سال ۱۳۹۸	۱۰۲
جدول ۵۴-۴	میانگین درصد خطا مطلق برای داده‌های پیش‌بینی سال ۹۸	۱۰۴
جدول ۵۵-۴	متوسط حجم ترافیک روزانه در سال ۱۳۹۸	۱۰۶
جدول ۵۶-۴	بهترین مدل پیش‌بینی در سه روش اعتبارسنجی RMSE، MAPE و AADT	۱۱۳

فهرست اشکال

شکل ۱-۲	الگوی پیش‌بینی و پیاده‌سازی مدل ARIMA به روش باکس جنکینز.....	۱۳
شکل ۲-۲	میزان رشد شبکه LSTM در گذر زمان.....	۱۵
شکل ۳-۲	میزان استفاده از یادگیری عمیق LSTM در روش‌های مختلف.....	۱۵
شکل ۴-۲	ساختار زنجیره‌ای در شبکه LSTM.....	۱۶
شکل ۵-۲	مقایسه تعداد ورودی و خروجی در شبکه RNN و LSTM.....	۱۶
شکل ۶-۲	مرحله نخست درک معماری LSTM.....	۱۷
شکل ۷-۲	گیت فراموشی شبکه LSTM.....	۱۷
شکل ۸-۲	مراحل تشکیل گیت فراموشی و ترکیب آن با C_{t-1}	۱۸
شکل ۹-۲	تشکیل اطلاعات استپ زمانی g_t	۱۹
شکل ۱۰-۲	افزافه شدن گیت ورودی به ساختار داخلی شبکه LSTM.....	۱۹
شکل ۱۱-۲	ساخت خروجی h_t	۲۰
شکل ۱۲-۲	نسخه کامل از شبکه عصبی LSTM همراه با تمامی فرمول‌های اثبات شده.....	۲۰
شکل ۱-۴	پلان مسیر مسیر مشهد - سرخس.....	۳۹
شکل ۲-۴	پلان مسیر مشهد - چناران.....	۳۹
شکل ۳-۴	پلان مسیر مشهد - فریمان.....	۴۰
شکل ۴-۴	پلان مسیر مشهد - ملک آباد.....	۴۰
شکل ۵-۴	پلان مسیر مشهد - سدکارده.....	۴۱
شکل ۶-۴	پلان مسیر مشهد - کلات.....	۴۱
شکل ۷-۴	حجم تردد روزانه مسیر مشهد - سرخس.....	۴۴
شکل ۸-۴	نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - سرخس.....	۴۵
شکل ۹-۴	حجم تردد روزانه مسیر سرخس - مشهد.....	۴۵
شکل ۱۰-۴	نمودار تغییرات حجم تردد نسبت به زمان در محور سرخس - مشهد.....	۴۶
شکل ۱۱-۴	حجم تردد روزانه مسیر مشهد - چناران.....	۴۷
شکل ۱۲-۴	نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - چناران.....	۴۸
شکل ۱۳-۴	حجم تردد روزانه مسیر چناران - مشهد.....	۴۸
شکل ۱۴-۴	نمودار تغییرات حجم تردد نسبت به زمان در محور چناران - مشهد.....	۴۹
شکل ۱۵-۴	حجم تردد روزانه مسیر مشهد - فریمان.....	۵۰
شکل ۱۶-۴	نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - فریمان.....	۵۱
شکل ۱۷-۴	حجم تردد روزانه مسیر فریمان - مشهد.....	۵۱
شکل ۱۸-۴	نمودار تغییرات حجم تردد نسبت به زمان در محور فریمان - مشهد.....	۵۲

- شکل ۴-۱۹ حجم تردد روزانه مسیر مشهد - ملک‌آباد ۵۳
- شکل ۴-۲۰ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - ملک‌آباد ۵۴
- شکل ۴-۲۱ حجم تردد روزانه مسیر ملک‌آباد - مشهد ۵۵
- شکل ۴-۲۲ نمودار تغییرات حجم تردد نسبت به زمان در محور ملک‌آباد - مشهد ۵۶
- شکل ۴-۲۳ حجم تردد روزانه مسیر مشهد - سدکارده ۵۶
- شکل ۴-۲۴ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - سدکارده ۵۷
- شکل ۴-۲۵ حجم تردد روزانه مسیر سدکارده - مشهد ۵۸
- شکل ۴-۲۶ نمودار تغییرات حجم تردد نسبت به زمان در محور سدکارده - مشهد ۵۹
- شکل ۴-۲۷ حجم تردد روزانه مسیر مشهد - کلات ۵۹
- شکل ۴-۲۸ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - سدکارده ۶۰
- شکل ۴-۲۹ حجم تردد روزانه مسیر کلات - مشهد ۶۱
- شکل ۴-۳۰ نمودار تغییرات حجم تردد نسبت به زمان در محور کلات - مشهد ۶۲
- شکل ۴-۳۱ نمونه‌ای از اجرای شبکه عصبی مصنوعی ۶۴
- شکل ۴-۳۲ تغییرات حجم تردد نسبت به زمان، پیش‌بینی و نمودار خطا در محور مشهد - سرخس ۶۴
- شکل ۴-۳۳ نمودارهای رگرسیون خطی برای داده‌های آموزش ۶۵
- شکل ۴-۳۴ نمودارهای رگرسیون خطی برای داده‌های آزمایش ۶۵
- شکل ۴-۳۵ تغییرات حجم تردد نسبت به زمان، پیش‌بینی و نمودار خطا در محور سرخس - مشهد ۶۶
- شکل ۴-۳۶ نمودارهای رگرسیون خطی برای داده‌های آموزش ۶۶
- شکل ۴-۳۷ نمودارهای رگرسیون خطی برای داده‌های آزمایش ۶۷
- شکل ۴-۳۸ تغییرات حجم تردد نسبت به زمان، پیش‌بینی و نمودار خطا در محور مشهد - چناران ۶۷
- شکل ۴-۳۹ نمودارهای رگرسیون خطی برای داده‌های آموزش ۶۸
- شکل ۴-۴۰ نمودارهای رگرسیون خطی برای داده‌های آموزش ۶۸
- شکل ۴-۴۱ تغییرات حجم تردد نسبت به زمان، پیش‌بینی و نمودار خطا در محور چناران - مشهد ۶۹
- شکل ۴-۴۲ نمودارهای رگرسیون خطی برای داده‌های آموزش ۶۹
- شکل ۴-۴۳ نمودارهای رگرسیون خطی برای داده‌های آزمایش ۷۰
- شکل ۴-۴۴ تغییرات حجم تردد نسبت به زمان، پیش‌بینی و نمودار خطا در محور مشهد - فریمان ۷۰
- شکل ۴-۴۵ نمودارهای رگرسیون خطی برای داده‌های آموزش ۷۱
- شکل ۴-۴۶ نمودارهای رگرسیون خطی برای داده‌های آزمایش ۷۱
- شکل ۴-۴۷ تغییرات حجم تردد نسبت به زمان، پیش‌بینی و نمودار خطا در محور فریمان - مشهد ۷۲
- شکل ۴-۴۸ نمودارهای رگرسیون خطی برای داده‌های آموزش ۷۲
- شکل ۴-۴۹ نمودارهای رگرسیون خطی برای داده‌های آزمایش ۷۳
- شکل ۴-۵۰ تغییرات حجم تردد نسبت به زمان، پیش‌بینی و نمودار خطا در محور مشهد - ملک‌آباد ۷۳

- شکل ۴-۵۱ نمودارهای رگرسیون خطی برای داده های آموزش ۷۴
- شکل ۴-۵۲ نمودارهای رگرسیون خطی برای داده های آزمایش ۷۴
- شکل ۴-۵۳ تغییرات حجم تردد نسبت به زمان، پیش بینی و نمودار خطا در محور ملک آباد - مشهد.. ۷۵
- شکل ۴-۵۴ نمودارهای رگرسیون خطی برای داده های آموزش ۷۵
- شکل ۴-۵۵ نمودارهای رگرسیون خطی برای داده های آزمایش ۷۶
- شکل ۴-۵۶ تغییرات حجم تردد نسبت به زمان، پیش بینی و نمودار خطا در محور مشهد - سدکارده .. ۷۶
- شکل ۴-۵۷ نمودارهای رگرسیون خطی برای داده های آموزش ۷۷
- شکل ۴-۵۸ نمودارهای رگرسیون خطی برای داده های آزمایش ۷۷
- شکل ۴-۵۹ تغییرات حجم تردد نسبت به زمان، پیش بینی و نمودار خطا در محور سدکارده - مشهد. ۷۸
- شکل ۴-۶۰ نمودارهای رگرسیون خطی برای داده های آموزش ۷۸
- شکل ۴-۶۱ نمودارهای رگرسیون خطی برای داده های آزمایش ۷۹
- شکل ۴-۶۲ تغییرات حجم تردد نسبت به زمان، پیش بینی و نمودار خطا در محور مشهد - کلات ۷۹
- شکل ۴-۶۳ نمودارهای رگرسیون خطی برای داده های آموزش ۸۰
- شکل ۴-۶۴ نمودارهای رگرسیون خطی برای داده های آزمایش ۸۰
- شکل ۴-۶۵ تغییرات حجم تردد نسبت به زمان، پیش بینی و نمودار خطا در محور کلات - مشهد ۸۱
- شکل ۴-۶۶ نمودارهای رگرسیون خطی برای داده های آزمایش ۸۱
- شکل ۴-۶۷ نمودارهای رگرسیون خطی برای داده های آزمایش ۸۲
- شکل ۴-۶۸ نمونه های از اجرای شبکه یادگیری عمیق LSTM ۸۳
- شکل ۴-۶۹ تغییرات حجم تردد نسبت به زمان، پیش بینی و نمودار خطا در محور مشهد - سرخس .. ۸۴
- شکل ۴-۷۰ نمودارهای رگرسیون خطی برای داده های آموزش ۸۴
- شکل ۴-۷۱ نمودارهای رگرسیون خطی برای داده های آزمایش ۸۴
- شکل ۴-۷۲ تغییرات حجم تردد نسبت به زمان، پیش بینی و نمودار خطا در محور سرخس - مشهد... ۸۵
- شکل ۴-۷۳ نمودار رگرسیون خطی داده های آموزش ۸۶
- شکل ۴-۷۴ نمودار رگرسیون خطی داده های آزمایش ۸۶
- شکل ۴-۷۵ تغییرات حجم تردد نسبت به زمان، پیش بینی و نمودار خطا در محور مشهد - چناران ۸۷
- شکل ۴-۷۶ نمودار رگرسیون خطی داده های آموزش ۸۷
- شکل ۴-۷۷ نمودار رگرسیون خطی داده های آزمایش ۸۷
- شکل ۴-۷۸ تغییرات حجم تردد نسبت به زمان، پیش بینی و نمودار خطا در محور چناران - مشهد ۸۸
- شکل ۴-۷۹ نمودار رگرسیون خطی داده های آموزش ۸۹
- شکل ۴-۸۰ نمودار رگرسیون خطی داده های آزمایش ۸۹
- شکل ۴-۸۱ تغییرات حجم تردد نسبت به زمان، پیش بینی و نمودار خطا در محور مشهد - فریمان ۹۰
- شکل ۴-۸۲ نمودار رگرسیون خطی داده های آموزش ۹۰

- شکل ۴-۸۳ نمودار رگرسیون خطی داده‌های آزمایش..... ۹۰
- شکل ۴-۸۴ تغییرات حجم تردد نسبت به زمان، پیش بینی و نمودار خطا در محور فریمان - مشهد.... ۹۱
- شکل ۴-۸۵ نمودار رگرسیون خطی داده‌های آموزش..... ۹۲
- شکل ۴-۸۶ نمودار رگرسیون خطی داده‌های آزمایش..... ۹۲
- شکل ۴-۸۷ تغییرات حجم تردد نسبت به زمان، پیش بینی و نمودار خطا در محور مشهد - ملک‌آباد. ۹۳
- شکل ۴-۸۸ نمودار رگرسیون خطی داده‌های آموزش..... ۹۳
- شکل ۴-۸۹ نمودار رگرسیون خطی داده‌های آزمایش..... ۹۳
- شکل ۴-۹۰ تغییرات حجم تردد نسبت به زمان، پیش بینی و نمودار خطا در محور ملک‌آباد - مشهد.. ۹۴
- شکل ۴-۹۱ نمودار رگرسیون خطی داده‌های آموزش..... ۹۵
- شکل ۴-۹۲ نمودار رگرسیون خطی داده‌های آزمایش..... ۹۵
- شکل ۴-۹۳ تغییرات حجم تردد نسبت به زمان، پیش بینی و نمودار خطا در محور مشهد - سدکارده.. ۹۶
- شکل ۴-۹۴ نمودار رگرسیون خطی داده‌های آموزش..... ۹۶
- شکل ۴-۹۵ نمودار رگرسیون خطی داده‌های آزمایش..... ۹۶
- شکل ۴-۹۶ تغییرات حجم تردد نسبت به زمان، پیش بینی و نمودار خطا در محور سدکارده - مشهد. ۹۷
- شکل ۴-۹۷ نمودار رگرسیون خطی داده‌های آموزش..... ۹۸
- شکل ۴-۹۸ نمودار رگرسیون خطی داده‌های آزمایش..... ۹۸
- شکل ۴-۹۹ تغییرات حجم تردد نسبت به زمان، پیش بینی و نمودار خطا در محور مشهد - کلات..... ۹۹
- شکل ۴-۱۰۰ نمودار رگرسیون خطی داده‌های آموزش..... ۹۹
- شکل ۴-۱۰۱ نمودار رگرسیون خطی داده‌های آزمایش..... ۹۹
- شکل ۴-۱۰۲ تغییرات حجم تردد نسبت به زمان، پیش بینی و نمودار خطا در محور کلات - مشهد... ۱۰۰
- شکل ۴-۱۰۳ نمودار رگرسیون خطی داده‌های آموزش..... ۱۰۱
- شکل ۴-۱۰۴ نمودار رگرسیون خطی داده‌های آزمایش..... ۱۰۱
- شکل ۴-۱۰۵ نمودار مقایسه خطای RMSE در مسیرهای موجود..... ۱۰۳
- شکل ۴-۱۰۶ نمودار مقایسه خطای RMSE در سه روش مدلسازی..... ۱۰۳
- شکل ۴-۱۰۷ مقایسه درصد میانگین خطای مطلق در سه روش مدلسازی..... ۱۰۵
- شکل ۴-۱۰۸ نمودار مقایسه خطای RMSE در سه روش مدلسازی..... ۱۰۵
- شکل ۴-۱۰۹ نتایج پیش بینی AADT در محور مشهد - سرخس..... ۱۰۷
- شکل ۴-۱۱۰ نتایج پیش بینی AADT در محور سرخس - مشهد..... ۱۰۷
- شکل ۴-۱۱۱ نتایج پیش بینی AADT در محور مشهد - چناران..... ۱۰۸
- شکل ۴-۱۱۲ نتایج پیش بینی AADT در محور سرخس - مشهد..... ۱۰۸
- شکل ۴-۱۱۳ نتایج پیش بینی AADT در محور مشهد - فریمان..... ۱۰۹
- شکل ۴-۱۱۴ نتایج پیش بینی AADT در محور فریمان - مشهد..... ۱۰۹

- شکل ۴-۱۱۵ نتایج پیش بینی AADT در محور مشهد - ملکآباد..... ۱۱۰
- شکل ۴-۱۱۶ نتایج پیش بینی AADT در محور ملکآباد - مشهد..... ۱۱۰
- شکل ۴-۱۱۷ نتایج پیش بینی AADT در محور مشهد - سدکارده..... ۱۱۱
- شکل ۴-۱۱۸ نتایج پیش بینی AADT در محور ملکآباد - مشهد..... ۱۱۱
- شکل ۴-۱۱۹ نتایج پیش بینی AADT در محور مشهد - کلات..... ۱۱۲
-
- شکل آ- ۱ نمودار ACF و PACF برای داده‌های اصلی در محور مشهد - سرخس..... ۱۲۳
- شکل آ- ۲ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور مشهد - سرخس..... ۱۲۳
- شکل آ- ۳ نمودارهای ACF و PACF باقیمانده‌های در محور مشهد - سرخس..... ۱۲۳
- شکل آ- ۴ نمودار احتمال نرمال باقیمانده‌ها در محور مشهد - سرخس..... ۱۲۴
- شکل آ- ۵ نمودار باقیمانده‌های مدل نسبت به زمان در محور مشهد - سرخس..... ۱۲۴
- شکل آ- ۶ نمودار ACF و PACF برای داده‌های اصلی در محور سرخس - مشهد..... ۱۲۴
- شکل آ- ۷ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور سرخس - مشهد..... ۱۲۵
- شکل آ- ۸ نمودارهای ACF و PACF باقیمانده‌های در محور سرخس - مشهد..... ۱۲۵
- شکل آ- ۹ نمودار احتمال نرمال باقیمانده‌ها در محور سرخس - مشهد..... ۱۲۵
- شکل آ- ۱۰ نمودار باقیمانده‌های مدل نسبت به زمان در محور سرخس - مشهد..... ۱۲۶
- شکل آ- ۱۱ نمودار ACF و PACF برای داده‌های اصلی در محور مشهد - چناران..... ۱۲۶
- شکل آ- ۱۲ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور مشهد - چناران..... ۱۲۶
- شکل آ- ۱۳ نمودارهای ACF و PACF باقیمانده‌های در محور مشهد - چناران..... ۱۲۷
- شکل آ- ۱۴ نمودار احتمال نرمال باقیمانده‌ها در محور مشهد - چناران..... ۱۲۷
- شکل آ- ۱۵ نمودار باقیمانده‌های مدل نسبت به زمان در محور مشهد - چناران..... ۱۲۷
- شکل آ- ۱۶ نمودار ACF و PACF برای داده‌های اصلی در محور چناران - مشهد..... ۱۲۸
- شکل آ- ۱۷ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور چناران - مشهد..... ۱۲۸
- شکل آ- ۱۸ نمودارهای ACF و PACF باقیمانده‌های در محور چناران - مشهد..... ۱۲۸
- شکل آ- ۱۹ نمودار احتمال نرمال باقیمانده‌ها در محور چناران - مشهد..... ۱۲۹
- شکل آ- ۲۰ نمودار باقیمانده‌های مدل نسبت به زمان در محور چناران - مشهد..... ۱۲۹
- شکل آ- ۲۱ نمودار ACF و PACF برای داده‌های اصلی در محور مشهد - فریمان..... ۱۲۹
- شکل آ- ۲۲ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور مشهد - فریمان..... ۱۳۰
- شکل آ- ۲۳ نمودارهای ACF و PACF باقیمانده‌های در محور مشهد - فریمان..... ۱۳۰
- شکل آ- ۲۴ نمودار احتمال نرمال باقیمانده‌ها در محور مشهد - فریمان..... ۱۳۰
- شکل آ- ۲۵ نمودار باقیمانده‌های مدل نسبت به زمان در محور مشهد - فریمان..... ۱۳۱
- شکل آ- ۲۶ نمودار ACF و PACF برای داده‌های اصلی در محور فریمان - مشهد..... ۱۳۱

شکل آ- ۲۷ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور فریمان - مشهد..... ۱۳۱
 شکل آ- ۲۸ نمودارهای ACF و PACF باقیمانده‌های در محور فریمان - مشهد..... ۱۳۲
 شکل آ- ۲۹ نمودار احتمال نرمال باقیمانده‌ها در محور فریمان - مشهد..... ۱۳۲
 شکل آ- ۳۰ نمودار باقیمانده‌های مدل نسبت به زمان در محور فریمان - مشهد..... ۱۳۲
 شکل آ- ۳۱ نمودار ACF و PACF برای داده‌های اصلی در محور مشهد - ملک‌آباد..... ۱۳۳
 شکل آ- ۳۲ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور مشهد - ملک‌آباد..... ۱۳۳
 شکل آ- ۳۳ نمودارهای ACF و PACF باقیمانده‌های در محور مشهد - ملک‌آباد..... ۱۳۳
 شکل آ- ۳۴ نمودار احتمال نرمال باقیمانده‌ها در محور مشهد - ملک‌آباد..... ۱۳۴
 شکل آ- ۳۵ نمودار باقیمانده‌های مدل نسبت به زمان در محور مشهد - ملک‌آباد..... ۱۳۴
 شکل آ- ۳۶ نمودار ACF و PACF برای داده‌های اصلی در محور ملک‌آباد - مشهد..... ۱۳۴
 شکل آ- ۳۷ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور ملک‌آباد - مشهد..... ۱۳۵
 شکل آ- ۳۸ نمودارهای ACF و PACF باقیمانده‌های در محور ملک‌آباد - مشهد..... ۱۳۵
 شکل آ- ۳۹ نمودار احتمال نرمال باقیمانده‌ها در محور ملک‌آباد - مشهد..... ۱۳۵
 شکل آ- ۴۰ نمودار باقیمانده‌های مدل نسبت به زمان در محور ملک‌آباد - مشهد..... ۱۳۶
 شکل آ- ۴۱ نمودار ACF و PACF برای داده‌های اصلی در محور مشهد - سدکارده..... ۱۳۶
 شکل آ- ۴۲ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور مشهد - سدکارده..... ۱۳۶
 شکل آ- ۴۳ نمودارهای ACF و PACF باقیمانده‌های در محور مشهد - سدکارده..... ۱۳۷
 شکل آ- ۴۴ نمودار احتمال نرمال باقیمانده‌ها در محور مشهد - سدکارده..... ۱۳۷
 شکل آ- ۴۵ نمودار باقیمانده‌های مدل نسبت به زمان در محور مشهد - سدکارده..... ۱۳۷
 شکل آ- ۴۶ نمودار ACF و PACF برای داده‌های اصلی در محور سدکارده - مشهد..... ۱۳۸
 شکل آ- ۴۷ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور سدکارده - مشهد..... ۱۳۸
 شکل آ- ۴۸ نمودارهای ACF و PACF باقیمانده‌های در محور سدکارده - مشهد..... ۱۳۸
 شکل آ- ۴۹ نمودار احتمال نرمال باقیمانده‌ها در محور سدکارده - مشهد..... ۱۳۹
 شکل آ- ۵۰ نمودار باقیمانده‌های مدل نسبت به زمان در محور سدکارده - مشهد..... ۱۳۹
 شکل آ- ۴۱ نمودار ACF و PACF برای داده‌های اصلی در محور مشهد - سدکارده..... ۱۳۹
 شکل آ- ۴۲ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور مشهد - سدکارده..... ۱۴۰
 شکل آ- ۴۳ نمودارهای ACF و PACF باقیمانده‌های در محور مشهد - سدکارده..... ۱۴۰
 شکل آ- ۴۴ نمودار احتمال نرمال باقیمانده‌ها در محور مشهد - سدکارده..... ۱۴۰
 شکل آ- ۴۵ نمودار باقیمانده‌های مدل نسبت به زمان در محور مشهد - سدکارده..... ۱۴۱
 شکل آ- ۴۶ نمودار ACF و PACF برای داده‌های اصلی در محور سدکارده - مشهد..... ۱۴۱
 شکل آ- ۴۷ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور سدکارده - مشهد..... ۱۴۱
 شکل آ- ۴۸ نمودارهای ACF و PACF باقیمانده‌های در محور سدکارده - مشهد..... ۱۴۲

- شکل آ- ۴۹ نمودار احتمال نرمال باقیمانده‌ها در محور صدکارده - مشهد..... ۱۴۲
- شکل آ- ۵۰ نمودار باقیمانده‌های مدل نسبت به زمان در محور صدکارده - مشهد..... ۱۴۲

فصل ۱: مقدمه

۱ - ۱ مقدمه

شبکه‌های اصلی ارتباطی مهمترین و حساس‌ترین ساختار صنعت حمل و نقل هر کشور محسوب می‌شود. ساخت و توسعه شبکه‌های ارتباطی، حفظ و نگهداری راه‌های موجود، تامین ایمنی و بالا بردن سطح خدمت‌رسانی در راه‌ها از مهمترین وظایف و اولویت‌های اصلی وزارت راه و ترابری می‌باشد که بدون پیش‌بینی و مدیریت درست و کارآمد محقق نمی‌شود.

تأثیرات منفی زیست محیطی، اجتماعی و اقتصادی که راه‌بندان ترافیکی و شلوغی راه‌ها بر جوامع بشری می‌گذارد محققین را برآن داشته است که به دنبال راه کارهایی برای کاهش شلوغی/ تاخیر و مقابله با آن باشند. این ازدحام منجر به افزایش زمان سفر، آلودگی، مصرف سوخت، کاهش بهره‌وری و نارضایتی مسافران می‌شود. از جمله این راه‌حل‌ها می‌توان به افزایش کیفیت حمل و نقل عمومی، استفاده از دوچرخه‌های کرایه‌ای و ایجاد کمربندی برای دور زدن شهرها و عدم ورود به ترافیک شهری اشاره کرد. از جمله روش‌های موجود برای رسیدن به این هدف، پیش‌بینی مسیر و مقصد نهایی خودرو و حجم تردد راه‌ها است. اگر حجم تردد راه‌ها را بتوان پیش‌بینی کرد، به راحتی می‌توان قادر به تخمین ازدحام ترافیک، سطح سرویس راه، شد. پیش‌بینی جریان ترافیک یکی از نیازهای اساسی سفر راحت است، اما این وظیفه به دلیل افزایش روزافزون ترافیک و عدم قطعیت داده‌های بزرگ ترافیک، چالش برانگیز است. با استفاده از داده‌های ترافیک و حمل و نقل که مشخصات جریان ترافیک و اطلاعات سفر ثبت شده‌اند و برای تجزیه و تحلیل در دسترس هستند.

۱ - ۲ تعریف مساله

اطلاعات دقیق و به موقع جریان ترافیک می‌تواند به کارشناسان حمل و نقل در برنامه‌ریزی زمان سفر و مسیر سفر کمک کند، که باعث کاهش ترافیک و کاهش سفرهای غیرضروری می‌شود. پیش‌بینی دقیق کوتاه مدت جریان ترافیک کلید ارائه چنین وضعیت ترافیکی دقیق و به موقع است.

پیش‌بینی جریان ترافیک به عنوان یک موضوع حیاتی و چالش برانگیز در دانشگاه و صنعت در نظر گرفته شده است. این پیش‌بینی به احتمال زیاد به کاربران جاده‌ای کمک می‌کند تا تصمیمات بهتری در سفر بگیرند، کارایی عملکرد ترافیک را افزایش دهند، انتشار کربن و ازدحام ترافیک را کاهش دهند. پیش‌بینی دقیق شرایط آینده ترافیک (به عنوان مثال، جریان ترافیک، سرعت سفر و زمان سفر) برای سیستم‌های حمل و نقل هوشمند (ITS)^۱ ضروری است، که می‌تواند به مدیران کمک کند تا اقدامات پیشگیرانه کافی را در برابر ازدحام ترافیک را انجام دهند و مسافران تصمیمات آگاهانه‌تری بگیرند. در بین کاربردهای مختلف در ITS، پیش‌بینی جریان ترافیک در چند دهه گذشته توجه زیادی را به خود جلب کرده است. جریان ترافیک هنوز یک موضوع چالش برانگیز برای محققان حمل و نقل است. با

^۱ Intelligent Transportation Systems (ITS)

توجه به ویژگی‌های تصادفی جریان ترافیک، پیش‌بینی دقیق ترافیک کار ساده‌ای نیست. برای حل این موضوع، تکنیک‌های زیادی برای مدلسازی و پیش‌بینی جریان ترافیک به کار گرفته شده‌است.

۱ - ۳ اهمیت پیش‌بینی و حوزه کاربرد

تحلیل سری زمانی^۱، در دهه‌های اخیر بسیاری از محققین را به سوی خود جذب کرده است. هدف اصلی در تحلیل سری زمانی در مورد یک پدیده، ایجاد یک مدل آماری برای داده‌های وابسته به زمان براساس اطلاعات گذشته آن پدیده است. با این کار امکان پیش‌بینی در مورد آینده پدیده مورد بحث میسر می‌شود. به بیان دیگر تحلیل سری زمانی، ایجاد مدلی گذشته‌نگر است تا امکان تصمیمات آینده‌نگر را فراهم سازد.

حجم تردد یکی از مهمترین ارکان اصلی ترافیک بشمار میرود که در تمامی مباحث حمل و نقل و ترافیکی ورودی مسایل می‌باشد. اطلاعات به دست آمده از حجم ترافیک در برنامه‌ریزی ترافیک، طراحی، عملکرد و تحقیقات بی‌اندازه مهم است و نوع اطلاعات مورد نیاز بسته به کاربرد آن می‌تواند تغییر کند.

۱- حجم کل ترافیک سالیانه بر حسب وسیله نقلیه در سال که در تعیین میزان مسافرت سالیانه بر حسب وسیله نقلیه در کیلومتر به عنوان معیار اقتصادی، اندازه‌گیری و مشخص کردن روند تغییرات حجم ترافیک و محاسبه میزان تصادف به نسبت حجم ترافیک در طول راه

۲- حجم ADT^2 (متوسط حجم ترافیک روزانه) و $AADT^3$ (متوسط حجم ترافیک روزانه در سال) بر حسب وسیله نقلیه در روز برای برنامه ریزی جاده‌ها (مانند ایجاد سیستم بزرگراه‌ها)، سیستم خیابان‌های اصلی و انتخاب بهترین مسیر برای توسعه یا ایجاد یک خیابان اصلی

۳- حجم‌های ساعتی بر حسب وسیله نقلیه در ساعت برای طرح هندسی در رابطه با تعیین تعداد و عرض خط‌های عبوری و جداسازی ترافیک، طرح تقاطع یا طرح ورودی‌ها و خروجی‌ها و برآورد کمبود ظرفیت

۴- حجم‌های کوتاه مدت با زمانهای کمتر از یک ساعت (۵، ۶، ۱۰ و ۱۵ دقیقه‌ای) برای تعیین تردد ماکزیمم، مشخص کردن تغییرات تردد در ساعت‌های اوج و تعیین محدودیت‌های ظرفیتی در مناطق شهری.

در پژوهش حاضر پیش‌بینی جریان ترافیک برای برنامه ریزی در جاده‌ها استفاده می‌گردد. بنابراین از اطلاعات نوع ۲ یعنی حجم ADT و $AADT$ بر حسب وسیله نقلیه در روز استفاده می‌گردد. موارد کاربرد حجم ترافیک جاده‌ای (نوع ۲) عبارتند از:

^۱ Time series

^۲ Average Daily Traffic

^۳ Average Annual Daily Traffic

- تعیین اهمیت جاده‌ها نسبت به یکدیگر
- اولویت بندی تعریض، تعمیر و نگهداری جاده‌ها
- ایجاد مسیر های جدید
- تعیین تغییرات تردد وسایل نقلیه در زمان‌های مختلف
- تعیین نحوه توزیع ترافیک در شبکه
- تعیین ظرفیت جاده‌ها
- مطالعه آثار ترافیم بر محیط زیست
- مطالعه تصادفات در جاده‌ها
- بررسی مسائل اقتصادی مربوط به ترافیک

۱ - ۴ اهداف تحقیق

آینده در هر حوزه‌ای، در دست آنهایی است که نگاهی رو به جلو در محور زمان دارند. تکنیک‌های پیش‌بینی برای وسایل نقلیه همیشه برای هرگونه ابتکار مهم و توسعه زیرساخت‌ها در بخش مهندسی حمل و نقل از اهمیت ویژه‌ای برخوردار بوده و خواهد بود. پیش‌بینی جریان ترافیک یک موضوع مهم تحقیقاتی برای جلوگیری از ازدحام ترافیک در سیستم‌های حمل و نقل است و جلوگیری از ازدحام ترافیک می‌تواند با شناخت درست از جریان ترافیک و سپس انجام برنامه‌ریزی حمل و نقلی به دست آید.

خراسان رضوی و مشهد به عنوان مقصد اصلی سفرهای جاده‌ای، سالانه دست‌کم ۲۰ میلیون مسافر می‌تواند شاخصی برای ارزیابی اقدامات حمل و نقل جاده‌های کشور در نظر گرفته شود. این استان با ۱۱۰ کیلومتر آزادراه در رتبه نهم کشور، ۷۲ کیلومتر بزرگراه در رتبه پنجم، یک هزار و ۲۸۹ کیلومتر راه اصلی در رتبه ششم، سه هزار و ۹۵۰ کیلومتر راه فرعی در رتبه چهارم در کشور دارا می‌باشد. همچنین این استان با بیست هزار و ۵۷۴ کیلومتر طول کل راه‌های برون شهری بعد از استان فارس در رتبه دوم قرار دارد.

ما یک رویکرد جدید پیش‌بینی جریان ترافیک را پیشنهاد کرده که وابستگی‌های زمانی و کوتاه‌مدت مشاهدات سری زمانی را به دست می‌آورد. با توجه به اهمیت پیش‌بینی درست حجم تردد در برنامه‌ریزی شبکه حمل و نقل جاده‌ای و توسعه آن، در این رساله پیش‌بینی و مدل سازی حجم تردد برای ۶ مسیر رفت و برگشت منتهی به مشهد انجام گرفته‌است. در تحقیق حاضر داده‌های موجود حاصل از داده‌های ارائه شده در سازمان راهداری و حمل و نقل جاده‌ای کشور می‌باشد.

این مسیرها عبارتند از :

- مشهد - سرخس
- مشهد - چناران
- مشهد - فریمان
- مشهد - ملک آباد
- مشهد - کلات
- مشهد - سدکارده

خصوصیات جاده‌های خارج شهری که در جاده‌های مختلف دیده می‌شود، مثلاً یک جاده تفریحی با ترافیک سنگین در تابستان و نقطه اوج در تعطیلی آخر هفته و یا یک جاده چند منظوره بین استانی بدون داشتن نقطه اوج که ممکن است در بعضی از فصول به خاطر سفرهای طولانی تفریحی بیشتر در آن فصل اوج داشته باشد.

۱ - ۵ فرضیه ها

همانطور که می‌دانیم حجم تردد در مسیرهای مختلف براساس نوع کاربرد مسیر و نوع سیستم کنترل ترافیک با هم تفاوت فراوانی دارد. در این رساله برای هر مسیر (رفت یا برگشت) مدلسازی‌های جداگانه‌ای صورت گرفته و مورد بررسی قرار گرفته‌است. همچنین از روش‌های پیش‌بینی معتبر و پرکاربرد در مسائل سری‌های زمانی استفاده شده است که در حمل و نقل و ترافیک توجه کمتری شده است. از سه روش پیش‌بینی ARIMA^۱ فصلی، شبکه عصبی مصنوعی (ANN^۲) و یادگیری عمیق LSTM^۳ بهره‌گرفته شده‌است. برای هر مسیر با توجه داده‌های موجود، داده‌های سال آخر (سال ۹۸) به منظور اعتبار سنجی به مدل انتخابی کنار گذاشته می‌شود و بهترین مدل برای هر مسیر در نظر گرفته می‌شود. در نهایت برای هر مسیر پیش‌بینی یکساله با درصد خطای بهینه انجام می‌شود.

^۱ AutoRegressive Integrated Moving Average (ARIMA)

^۲ Artificial neural network (ANN)

^۳ Long short-term memory (LSTM)

۱ - ۶ روش و مراحل تحقیق

۱ - ۶ - ۱ روش ARIMA

هدف اصلی در تحلیل سری زمانی، ایجاد یک مدل آماری برای داده‌های وابسته به زمان براساس اطلاعات گذشته آن پدیده است. با این کار امکان پیش‌بینی در مورد آینده پدیده مورد بحث میسر می‌شود. فرایند پیش‌بینی در سری زمانی شامل ۱- تشخیص ایستایی^۱ یا نایستایی^۲ سری زمانی از طریق رسم نمودار داده‌ها، ۲- هموارسازی^۳ سری زمانی نایستا ۳- عمل پیش‌بینی و تشخیص نوع مدل ۴- برازش مدل ۵- نکویی برازش مدل ۶- پیش‌بینی نهایی ۷- اعتبارسنجی مدل است. پیش‌بینی برای سری‌های زمانی نایستا به راحتی امکان پذیر نیست پس بهتر است عواملی که باعث خارج شدن سری زمانی از حالت ایستایی هستند، حذف شوند. به این ترتیب باید مولفه‌های شناسایی شده در سری زمانی حذف شوند، به این کار هموارسازی می‌گویند. روش‌های مختلفی برای هموارسازی سری زمانی وجود دارد از جمله عملگرهای میانگین متحرک^۴، هموارسازی نمایی^۵ ساده، روش‌های تفاضل‌گیری^۶ و ... که به حذف مولفه‌های سری زمانی کمک می‌کنند. همچنین می‌توان از نمودارهای مربوط به داده‌های سری زمانی برحسب زمان ترسیم می‌شوند مولفه‌های سری زمانی روند (Trend)، تناوب (Cyclic)، فصل (Seasonal) و تغییرات نامعمول (Irregular) را تشخیص داده و در نتیجه شناخت بهتری از داده‌های سری زمانی داشت.

۱ - ۶ - ۲ روش شبکه عصبی و LSTM

شبکه‌های عصبی از مجموعه نودها (به عنوان نرون^۷) و یال‌ها (در نقش سیناپس^۸) برای پردازش داده بهره می‌گیرند. در این سیستم، ورودی‌ها در شبکه به جریان افتاده و یک سری خروجی تولید می‌گردد. پس خروجی‌ها با داده‌های معتبر مقایسه می‌گردند. شبکه‌های عصبی اطلاعات را به روشی مشابه با کاری که مغز انسان انجام می‌دهد پردازش می‌کنند. آن‌ها از تعداد زیادی از عناصر پردازشی (سلول عصبی) که فوق العاده بهم پیوسته‌اند تشکیل شده است که این عناصر به صورت موازی باهم برای حل یک مسئله مشخص کار می‌کنند. شبکه‌های عصبی با مثال کار می‌کنند و نمی‌توان آن‌ها را برای انجام

^۱ Stationary

^۲ Non-stationary

^۳ Smoothing

^۴ Moving Average

^۵ Exponential Smoothing

^۶ Difference

^۷ Neurons

^۸ Synapses

یک وظیفه خاص برنامه ریزی کرد. مثال‌ها می‌بایست با دقت انتخاب شوند در غیر این صورت باعث اتلاف وقت و هزینه می‌شود و حتی بدتر از آن، ممکن است شبکه درست کار نکند.

از مزیت‌های شبکه عصبی می‌توان به موارد زیر اشاره نمود:

- **یادگیری انطباق پذیر^۱**: یادگیری انطباق پذیر، قابلیت یادگیری و نحوه انجام وظایف بر پایه اطلاعات داده شده برای تمرین و تجربه‌های مقدماتی.

- **سازماندهی توسط خود^۲**: سازماندهی توسط خود یعنی یک شبکه هوش مصنوعی سازماندهی یا ارائه‌اش را برای اطلاعاتی که در طول دوره یادگیری دریافت می‌کند، خودش ایجاد کند.

- **عملکرد بهنگام^۳**: در عملکرد بهنگام محاسبات شبکه هوش مصنوعی می‌تواند بصورت موازی انجام شود و سخت افزارهای مخصوصی طراحی و ساخته شده که می‌تواند از این قابلیت استفاده کنند.

شبکه‌ی LSTM سازوکارهایی داخلی به اسم گیت^۴ دارد. این گیت‌ها جریان اطلاعات را کنترل می‌کنند؛ همین‌طور مشخص می‌کنند چه داده‌هایی در توالی مهم هستند و باید هم‌چنان حفظ بشوند و چه داده‌هایی باید حذف بشوند؛ به این شکل، شبکه‌ی اطلاعات مهم را در طول زنجیره‌ی توالی عبور می‌دهد تا خروجی مدنظر را داشته باشیم.

در پژوهش حاضر، دو مدل شبکه عصبی مصنوعی و یادگیری عمیق LSTM کاملاً مشابه با یکدیگر کدنویسی شده‌اند. فقط با این تفاوت که نوع آموزش هر شبکه با هم متفاوت می‌باشد. از آنجایی که داده‌های سری زمانی وابسته به زمان بر اساس اطلاعات گذشته می‌باشند در اولین قدم به وابستگی داده‌ها نسبت به هم پرداخته می‌شوند بدین صورت که داده‌های هر روز به کدامیک از داده‌های روزهای قبلاز خود وابسته می‌باشد و در نهایت منجر به تشکیل یک ماتریس می‌شود که برای آموزش وارد شبکه می‌شود و عمل پیش‌بینی و مدلسازی انجام می‌گیرد.

۱ - ۷ ضرورت حل مسئله

در سال‌های اخیر شاهد ظهور دو فناوری مهم، سیستم‌های داده بزرگ که قادر به ذخیره، بازیابی و پردازش حجم زیادی از داده‌ها هستند و الگوریتم‌های یادگیری ماشین که قادر به یادگیری و پیش‌بینی توالی‌های پیچیده هستند. در ترکیب، این فناوری‌ها فرصت‌های جدیدی را برای استفاده از حجم زیادی

^۱ Adaptive Learning

^۲ Self Organization

^۳ Real Time Operation

^۴ Gate

از داده‌های حجم ترافیک برای بهبود پیش بینی جریان ترافیک و تشخیص جریان‌های غیر خطی ترافیک ارائه می دهند. از این رو در پژوهش حاضر از روش‌های پارامتریک (ARIMA)، غیرپارامتریک (ANN) و روش یادگیری عمیق LSTM که نه تنها وابستگی‌های زمانی و کوتاه مدت مشاهدات سری زمانی را در نظر می گیرد بلکه الگوهای مفقود شده برای بهبود نتایج پیش بینی را ارائه می دهد، استفاده شده است.

۱ - ۸ چالش‌های تحقیق

- مهمترین قدم در روش ARIMA چگونگی هموارسازی سری زمانی نایستا می باشد.
- چگونگی برازش مدل و ارائه پارامترهای مجهول مدل ARIMA.
- در برخی موارد به دلایل مختلف داده ها در دسترس نیستند چگونه داده ها را با کمترین خطا به دست آورد.
- پشتوانه تئوری ضعیف در روش LSTM به دلیل سال‌های اولیه و شبکه‌های موجود به صورت آزمون و خطا صورت می گیرد.
- آیا وابستگی‌های تعریف شده در روش شبکه عصبی مصنوعی و یادگیری عمیق LSTM برای هر روز (به چه روزهای قبل از خود وابسته و تاثیر می پذیرد) مناسب می باشد.
- دستیابی به کمترین میزان خطا ایجاد شده در هر مسیر و هر روش پیش‌بینی.
- مشخص کردن بهترین و نوین‌ترین روش پیش‌بینی روزانه برای داده‌های بزرگ حجم تردد جاده‌ای

۱ - ۹ ساختار پژوهش

در پژوهش انجام شده به ارائه مدل‌های پیش‌بینی حجم تردد ترافیک با استفاده از روش‌های ARIMA، شبکه عصبی مصنوعی و یادگیری عمیق LSTM پرداخته می شود. پژوهش حاضر در پنج فصل تدوین و ارائه می گردد. فصل اول شامل مقدمه‌ای از کلیات تحقیق و اهداف تحقیق می باشد. در فصل دوم به بیان مبانی نظری و پیشینه تحقیق پرداخته می شود. در فصل سوم روش‌ها و مدل‌های مورد استفاده شرح داده می شوند. در فصل چهارم به پیش‌بینی و مدلسازی با هریک از روش‌های ارائه شده پرداخته می شود و نتایج حاصل از هر روش مورد بحث و بررسی قرار می گیرد. در نهایت در فصل پنجم بهترین روش پیش‌بینی حجم تردد ترافیک و نیز پیشنهادات ارائه می گردد.

فصل ۲ : مبانی نظری و پیشینه تحقیق

۲ - ۱ مقدمه

از کاربردهای مهم تحلیل سری‌های زمانی می‌توان به پیش‌بینی رفتار پدیده‌های تصادفی مرتبط با زمان اشاره کرد و به وسیله مدل‌سازی و بررسی سازگاری داده‌ها با یکی از مدل‌های مختلف سری زمانی، قادر به تعیین مقدارهای آتی پدیده تصادفی شد.

۲ - ۲ مبانی نظری

۲ - ۲ - ۱ سری‌های زمانی

دنباله‌ای از داده‌ها که در یک محدود زمانی جمع‌آوری شده‌اند، یک سری زمانی را تشکیل می‌دهند. این داده‌ها تغییراتی که پدیده در طول زمان دچار شده را منعکس می‌کنند. بنابراین می‌توانیم این مقادارها را یک بردار وابسته به زمان بدانیم. در این حالت اگر X یک بردار باشد، سری زمانی را می‌توان به صورت زیر نشان داد؛ که در آن t ، بیانگر زمان و X نیز یک متغیر تصادفی است.

$$X(t), t=0, 1, 2, \dots$$

تعریف ۱-۲

طبق این تعریف زمان $t=0$ نیز قابل تعریف است. این لحظه می‌تواند زمان تولد یک پدیده یا هنگامی باشد که اولین اطلاعات در آن لحظه ثبت شده است. به این ترتیب $X(t)$ متغیر تصادفی X را در زمان t نشان می‌دهد. مقدارهای مشاهده شده این متغیر تصادفی دارای ترتیبی هستند که زمان وقوع هر داده را نشان می‌دهند.

اگر متغیر تصادفی X ، یک بعدی باشد، یعنی از بین ویژگی‌های مختلف یک پدیده فقط از یکی ویژگی برای ایجاد مدل سری‌زمانی استفاده شود، مدل یک متغیره^۱ نامیده می‌شود. ولی اگر از چندین ویژگی برای ایجاد مدل سری‌زمانی استفاده شود، مدل سری زمانی چند متغیره^۲ نامیده می‌شود. البته اگر علاوه بر زمان، مکان یا مختصات را (یا هر اطلاعاتی که مقدار داده‌ها به آن وابسته باشند) به مدل اضافه شوند، مربوط مبحث آمار فضایی^۳ می‌شود.

همچنین اگر تغییرات پدیده در مدل سری‌زمانی برای زمان‌های منقطع در نظر گرفته شود، سری را زمان-گسسته^۴ و برعکس اگر زمان به صورت پیوسته در مدل فرض شود، سری را زمان-پیوسته^۵ نامیده می‌شود.

^۱ Univariate

^۲ Multivariate

^۳ Spatial Statistic

^۴ Discreet Time

^۵ Continuous Time

معمولاً در سری زمان-گسسته، داده‌ها در مقاطع مشخصی از زمان مثل ساعت، روز یا هفته و حتی سال جمع‌آوری می‌شوند. غالباً ایجاد مدل‌ها برای سری‌های زمان-گسسته انجام می‌شود زیرا با استفاده از گروه‌بندی و ایجاد فاصله‌های زمانی ترتیبی، امکان تبدیل سری‌های زمانی-پیوسته به زمان-گسسته وجود دارد.

۲-۱-۲-۱ مولفه‌های یک سری زمانی

معمولاً می‌توان الگوی رفتار یا مدل تغییرات یک سری زمانی را به چهار مولفه تفکیک کرد. روند، تناوب، فصل و تغییرات نامعمول. اگر نمودار مربوط به داده‌های سری زمانی را برحسب زمان ترسیم شوند می‌توان این مولفه‌ها را تشخیص داد. در نتیجه شناخت بهتری از داده‌های سری زمانی داشت.

- روند: تمایل سری زمانی به افزایش، کاهش یا حتی ثابت بودن، روند را تشکیل می‌دهد. در یک سری زمانی با روند افزایشی، انتظار داریم مقدارهای سری زمانی در زمان‌های $t=1$ و $t=2$ به صورت $X(1) \leq X(2)$ باشند.
- تناوب: تغییرات یکسان و تکراری در مقاطع میان‌مدت، تناوب در سری زمانی نامیده می‌شود.
- فصل: در سری زمانی، تغییراتی که در دوره‌ای کوتاه‌تر از یک تناوب به صورت تکراری رخ می‌دهد، به تغییرات فصلی معروف است.
- تغییرات نامعمول: این گونه تغییرات بر اثر عوامل تصادفی و غیرقابل پیش‌بینی ایجاد می‌شوند.

۲-۲-۲ روش ARIMA

از شیوه و مدل‌های مختلفی برای تعیین ساختار داده‌های حاصل از یک پدیده وابسته به زمان استفاده می‌شود. یکی از کامل‌ترین مدل‌ها را می‌توان مدل SARIMA فصلی در نظر گرفت که اغلب به آن مدل خودهمبسته یک‌پارچه میانگین متحرک فصلی می‌گویند. چنین مدلی از چهار بخش تشکیل شده است. بخش اول، خودهمبسته^۱ بودن داده‌ها را مشخص و مدل‌بندی می‌کند. بخش دوم یک‌پارچگی^۲، بخش سوم میانگین متحرک^۳ است که برای هموارسازی مقادیر سری زمانی و بخش چهارم برای تغییرات فصلی^۴ به کار می‌رود. ترکیب این مدل‌ها، منجر به وجود مدلی کامل و توانا خواهد شد. چنین مدلی براساس این پارامترها به صورت تعریف ۱-۲ نمایش داده می‌شود:

$$ARIMA(p,d,q)(P,D,Q,s)$$

تعریف ۲-۲

^۱ Autoregressive

^۲ Integrated

^۳ Moving Average

^۴ Seasonality

پارامترهای موجود در بخش اول و دوم این مدل، به ترتیب مربوط به مولفه های تغییرات غیرفصلی و فصلی می باشد. پارامترهای مولفه های غیرفصلی به صورت p, d, q است که به ترتیب آن ها را با p برای مدل خودهمبستگی (AR)، پارامتر d یا درجه تفاضلی برای بخش یک پارچگی (Integrated)، پارامتر q برای مدل میانگین متحرک (MA) و پارامترهای مولفه های فصلی به صورت P, D, Q, s هستند s طول فصل ها را مشخص می کند. از طرفی پارامترهای P و Q نیز مانند p و q در مدل MA و AR با توجه به تغییرات فصلی تعیین می شوند. در انتها نیز پارامتر D به عنوان پارامتر یک پارچه سازی فصلی برای حذف تغییرات فصلی به مدل اضافه می شود را نشان می دهند، مشخصه های اصلی مدل ARIMA محسوب می شوند.

برای تخمین پارامترهای یک مدل ARIMA فصلی می توان از روش باکس جنکینز^۱ استفاده کرد. در این روش مدل ARIMA فصلی در رابطه ۱-۲ فرموله شده است.

$$\text{رابطه ۱-۲} \quad Y_t = (1 - \Theta_1 B - \dots - \Theta_p B^p) (1 - \Phi_1 B^s - \dots - \Phi_P B^{Ps}) (1 - B)^d (1 - B^s)^D Y_t =$$

$$(1 - \theta_1 B - \dots - \theta_q B^q) (1 - \Theta_1 B^s - \dots - \Theta_Q B^{Qs}) \text{ et}$$

در رابطه ۱-۲ به ترتیب پارامترهای $\Theta, \Phi, \theta, \Theta$ ضرایب محاسبه شده برای مولفه های اتورگرسیو غیرفصلی و فصلی و مولفه های میانگین متحرک غیر فصلی و فصلی می باشند. همچنین B (عملگر پسرو) یک ابزار نمادین سودمند در روش باکس جنکینز می باشد و به صورت رابطه ۲-۲ تعریف می شود.

$$\text{رابطه ۲-۲} \quad B^n Y_t = Y_{t-n}$$

جهت پیش بینی و پیاده سازی مدل ARIMA به روش باکس جنکینز، با توجه به شکل ۱-۲ سه فاز اصلی، تشخیص مدل، برآورد و آزمون مدل و کاربرد مدل تشکیل شده است.

فاز ۱: بعد از آماده سازی داده های اصلی، نمودار زمانی داده ها رسم می گردد و در صورت نایست بودن داده ها، از طریق تفاضلی کردن داده ها سری را ایستا می نماییم. در مرحله بعد از نمودارهای خودهمبستگی ACF^۲ و خودهمبستگی جزئی PACF^۳ برای شناسایی مدل اولیه استفاده می شود.

فاز ۲: این فاز مهمترین قسمت برای پیش بینی و مدلسازی می باشد. در مرحله اول پارامترهای مدل اولیه برآورد می شود و ادامه بهترین مدل انتخاب می شود. در مرحله دوم به تشخیص مدل انتخابی پرداخته می شود. برای این منظور مقادیر باقیمانده هریک از داده ها به دست آورده می شود و به بررسی نمودارها و آزمون های مربوط به آن ها پرداخته می شود:

^۱ Box-Jenkins

^۲ Auto Correlation Function (ACF)

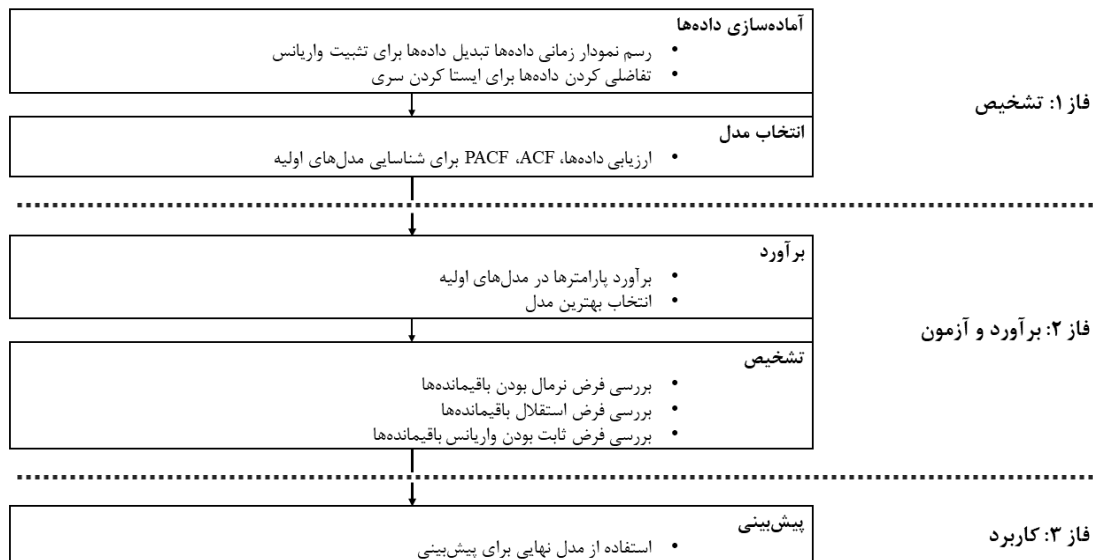
^۳ Partial Auto Correlation Function (PACF)

- بررسی فرض نرمال بودن باقیمانده‌ها

- بررسی فرض استقلال باقیمانده‌ها

- بررسی فرض ثابت بودن واریانس باقیمانده‌ها

فاز ۳: در صورت مناسب بودن مدل به پیش‌بینی نهایی پرداخته می‌شود.



شکل ۱-۲ الگوی پیش‌بینی و پیاده‌سازی مدل ARIMA به روش باکس جنکینز

۲ - ۲ - ۳ مدل شبکه عصبی مصنوعی و مدل LSTM

۲-۳-۱ یادگیری عمیق

یادگیری عمیق زیر مجموعه‌ای از یادگیری ماشین است در حالی که یادگیری عمیق و یادگیری ماشین هر دو زیرمجموعه‌ای از هوش مصنوعی می‌باشند. پس یادگیری عمیق زیر شاخه‌ای از یادگیری ماشین و هوش مصنوعی می‌باشد. یادگیری ماشین دارای سه مرحله مشترک می‌باشد:

- اطلاعات موجود از موضوع (داده‌ها)

- آموزش الگوریتم برای رسیدن به هدف مورد نظر (یادگیری)

- استفاده در کاربرد مورد نظر

مهم‌ترین بخش در یادگیری ماشین مرحله دوم (یادگیری) می‌باشد که در این مرحله به استخراج و بازنمایی ویژگی‌های مفید داده‌ها و در نهایت به آموزش الگوریتم‌ها پرداخته می‌شود. استخراج ویژگی‌ها

در مرحله یادگیری به دو صورت دستی و اتوماتیک انجام می‌شود که استخراج ویژگی‌ها به صورت اتوماتیک مربوط به مباحث یادگیری عمیق می‌باشد.

۲-۳-۲-۲ تاریخچه یادگیری عمیق

در سال ۱۹۹۳ میلادی اولین نمونه موفق در یادگیری عمیق، شبکه Lenet بود که کار این شبکه خواندن اعداد و کاراکترهای دست‌نویس بود که برای چک‌ها در آمریکا مورد استفاده قرار گرفت. چند سال بعد در سال ۱۹۹۷ مدل‌های RNN و LSTM ابداع شد. بعد از این دوره یادگیری عمیق به دلیل کمبود داده‌ها و سخت افزارهای مورد نیاز و همچنین مشکلات آموزش دچار یک رکود شد که این رکود تا سال ۲۰۰۶ ادامه داشت. اما در سال ۲۰۰۶ با مقاله آقای همینگتن یک روش جدید آموزش شبکه‌های عصبی تحت عنوان پیش‌آموزش لایه به لایه معرفی شد که خیلی از کمبودهای این روش را برطرف نمود. اوج محبوبیت این روش در سال ۲۰۱۲ اتفاق افتاد که توانست مقام اول رقابت‌های شناسایی کاراکترها، علایم ترافیکی، تصاویر پزشکی و Imagenet را به دست آورد.

۳-۳-۲-۲ مزایا یادگیری عمیق

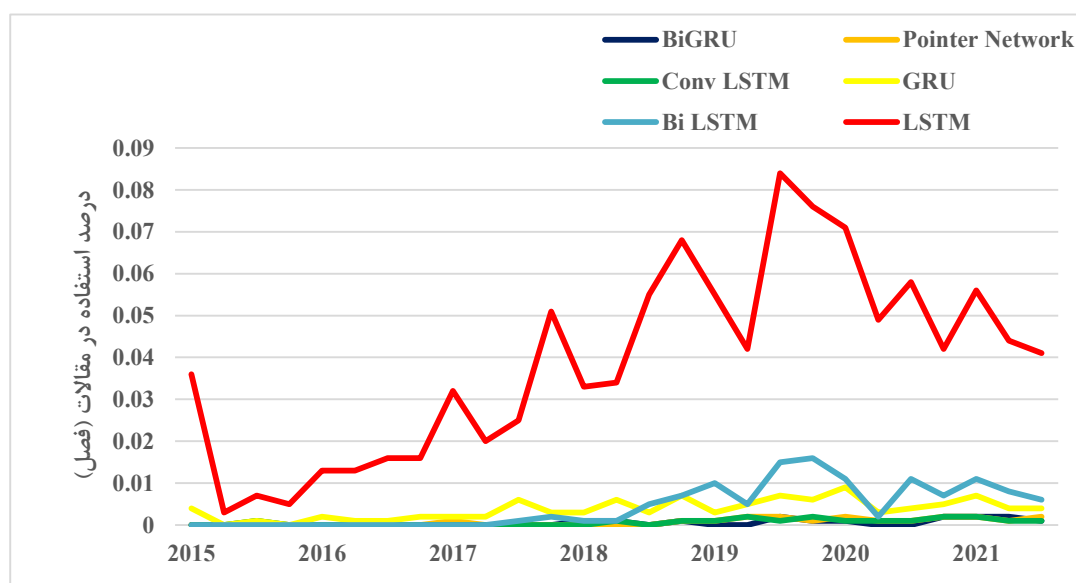
- یادگیری خودکار ویژگی‌ها
- یادگیری چند لایه ویژگی‌ها
- دقت بالا
- قدرت تعمیم بالا
- پشتیبانی سخت افزاری و نرم‌افزاری بالا
- پتانسیل ایجاد قابلیت‌های بیشتر

۴-۳-۲-۲ چالش‌های یادگیری عمیق

- پشته‌توانه تئوری ضعیف (سال‌های اولیه و شبکه‌های موجود به صورت آزمون و خطا صورت می‌گیرد)
- هزینه محاسباتی بالا
- نیاز به تعداد زیادی داده
- دشواری تنظیم پارامترها
- مشکلات آموزش

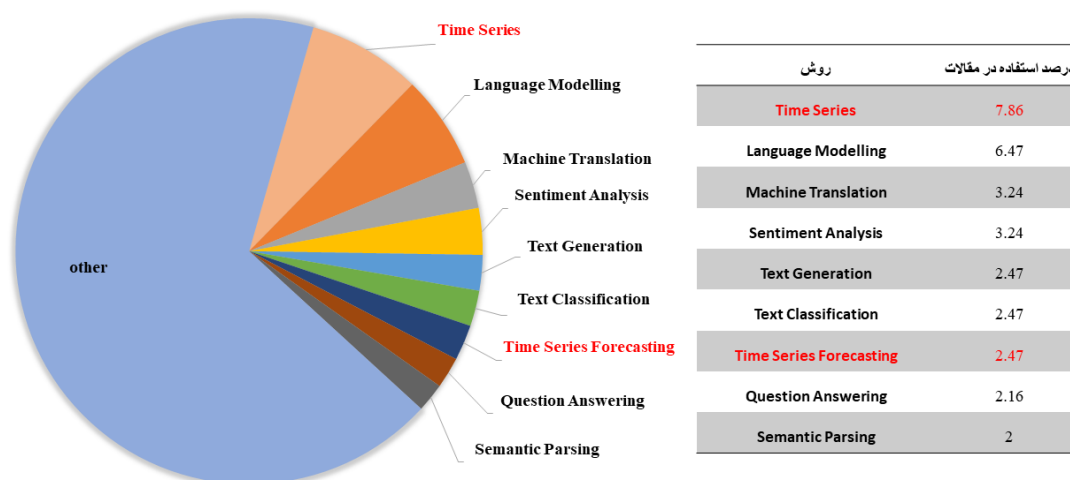
۵-۳-۲-۲ مدل LSTM

شبکه عصبی LSTM اولین بار توسط Hochreiter و Schmidhuber در سال ۱۹۹۷ معرفی شد. البته، در سال‌های بعد این کار توسط بسیاری از افراد بهبود داده شد. شبکه LSTM یک شبکه نسبتاً قدیمی است. اما در طیف وسیعی از مسائل استفاده می‌شود و هنوز از محبوبیت بالایی برخوردار است. با توجه به شکل ۲-۲ مشاهده می‌شود که با گذشت زمان، میزان استفاده از LSTM به شدت رشد داشته است.



شکل ۲-۲ میزان رشد شبکه LSTM در گذر زمان

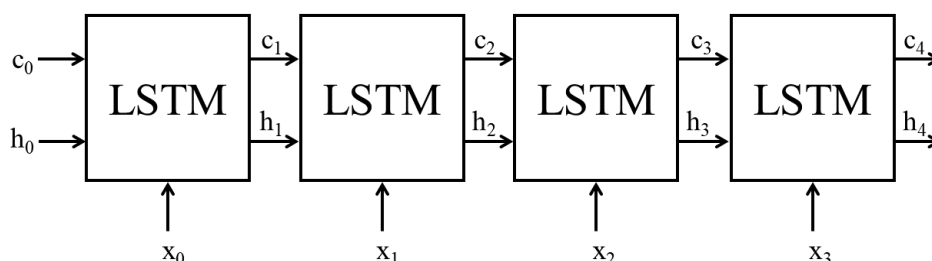
همچنین نمودار شکل ۳-۲ میزان استفاده LSTM در کاربردهای مختلف اشاره دارد. Other اشاره به روش‌های مختلف و کمتر شناخته شده دارد. Other را کنار بگذاریم، می‌بینیم که بیشترین میزان استفاده از LSTM در داده‌های سری زمانی است.



شکل ۳-۲ میزان استفاده از یادگیری عمیق LSTM در روش‌های مختلف

۲-۳-۶ ساختار کلی شبکه عصبی LSTM

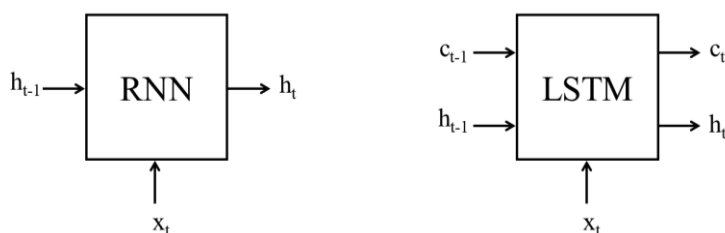
شبکه عصبی LSTM همانند شبکه RNN به صورت زنجیره‌ای پشت سرهم قرار می‌گیرد. شکل زیر این ساختار زنجیره‌ای را نشان می‌دهد.



شکل ۲-۴ ساختار زنجیره‌ای در شبکه LSTM.

لزوم وجود شبکه عصبی بازگشتی نسبت به سایر شبکه‌های عصبی در این است که این نوع شبکه‌ها برای پردازش داده‌های سری زمانی و ترتیبی مناسب است. همچنین برای شبکه بازگشتی مولفه زمان از اهمیت زیادی برخوردار است.

شبکه LSTM نقاط قوتی نسبت به شبکه RNN دارد و LSTM برای حل ضعف‌های RNN پیشنهاد شده است. ضعف‌های قابل توجهی که نمی‌توان آن‌ها را نادیده گرفت. شبکه عصبی RNN مشکل وابستگی بلندمدت یا Long Term Dependency دارد که شبکه LSTM این مشکل را برطرف می‌نماید و دلیل این امر مربوط به معماری و ساختار LSTM می‌باشد. همان‌طور که در شکل ۲-۵ مشخص می‌باشد شبکه LSTM نسبت به RNN دو ورودی و دو خروجی دارد.



شکل ۲-۵ مقایسه تعداد ورودی و خروجی در شبکه RNN و LSTM

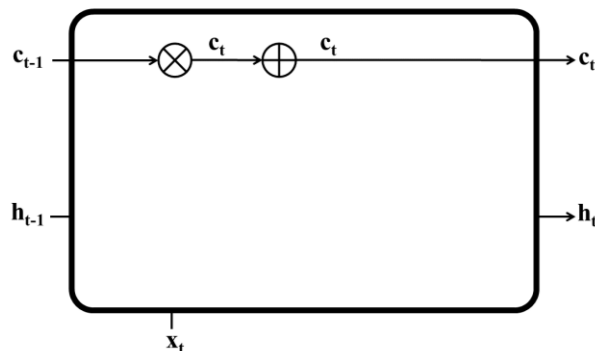
۲-۳-۷ معماری یادگیری عمیق LSTM

حال مرحله به مرحله معماری و ساختار شبکه LSTM به صورت دقیق مورد بررسی قرار گرفته است.

مرحله اول: بین این ورودی و خروجی‌ها، یکی از ورودی‌ها مستقیم به خروجی متصل می‌شود. در شکل ۲-۶ ورودی C_{t-1} مستقیماً به خروجی C_t متصل شده است. این اتصال همین‌طور ساده از اول تا آخر دنباله ادامه دارد. C مخفف Cell State هست و یک مولفه کلیدی در LSTM است. به Cell State، حافظه بلندمدت یا Long Term Memory هم گفته می‌شود. خط ورودی-خروجی از ابتدا تا

انتها، همه بلوک‌های LSTM در استپ‌های زمانی مختلف را به هم متصل می‌کند. در گذر زمان، در این حافظه بلندمدت (C) اطلاعاتی ذخیره یا از آن حذف می‌شود. این حافظه بلندمدت دو خاصیت مهم و جالب دارد:

- می‌تواند اطلاعات در آن را پاک کند. یعنی، همان فراموش کردن.
- می‌تواند به آن اطلاعاتی اضافه کند. یعنی، همان به‌خاطر سپردن.



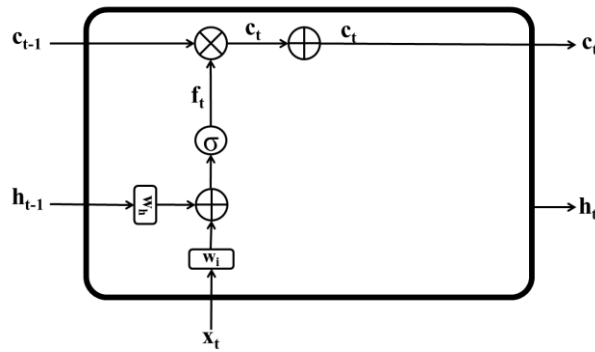
شکل ۶-۲ مرحله نخست درک معماری LSTM

مرحله دوم: در این مرحله مفهوم گیت فراموشی (forget gate) مشخص می‌شود. این گیت به صورت شکل ۷-۲ می‌باشد. یک شبکه عصبی کوچک با دو ورودی x_t و h_{t-1} تشکیل می‌دهد. این شبکه عصبی کوچک، وظیفه فراموشی بخشی از اطلاعات موجود در حافظه بلندمدت را دارد. به این شبکه عصبی گیت فراموشی گفته می‌شود. با گیت فراموشی، شبکه LSTM می‌تواند تنظیم کند که چه چیزی را فراموش کند و البته، چه چیزی را همچنان در حافظه نگه دارد.



شکل ۷-۲ گیت فراموشی شبکه LSTM

مرحله سوم: در این مرحله به جزئیات گیت فراموشی و ترکیب آن با C_{t-1} پرداخته می‌شود. این شبکه دو ورودی x_t و h_{t-1} دارد. این دو ورودی با هم ترکیب می‌شوند و سپس از یک لایه سیگموئید (σ) می‌گذرند. لایه سیگموئید عددی بین ۰ تا ۱ ایجاد می‌کند که در بردار ورودی C_{t-1} ضرب درایه می‌شود. f_t یک بردار است نه یک عدد اسکالر تنها. این بردار به اندازه C_{t-1} است که تعیین می‌کند که هر درایه C_{t-1} باید در چه عددی از ۰ تا ۱ ضرب شود. هر درایه از f_t اگر به عدد ۱ نزدیک باشد، به معنی این است که این درایه از C_{t-1} باید حفظ شود و اگر به ۰ نزدیک باشد، یعنی این درایه از C_{t-1} را باید دور ریخته شود (شکل ۸-۲).



شکل ۸-۲ مراحل تشکیل گیت فراموشی و ترکیب آن با C_{t-1}

در فرمول اثبات ۱-۲، نحوه به دست آوردن f_t را نمایش می‌دهد:

$$f_t = \sigma (w_{hf} \cdot h_{t-1} + w_{if} \cdot x_t + b_{hf} + b_{if}) \quad \text{اثبات ۱-۲}$$

در اثبات ۱-۲ دو لایه فولیکانکتد با وزن‌های w_{hf} و w_{if} وجود دارد. این دولایه به ترتیب برای ورودی h_{t-1} و x_t هستند. اندیس hf مخفف عبارت hidden و forget که بیانگر ورودی h_{t-1} و مربوط به بخش گیت فراموشی می باشد. اندیس if مخفف عبارت input و forget که بیانگر ورودی x و مربوط به بخش گیت فراموشی می باشد.

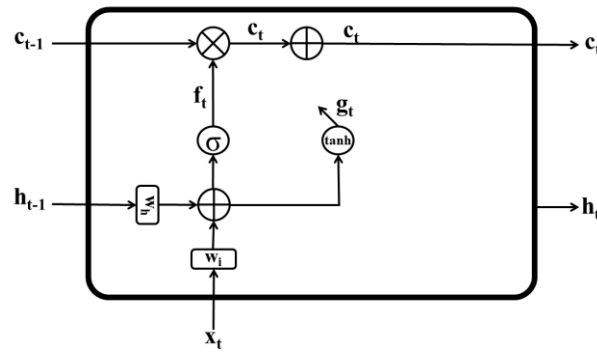
بعد از به دست آوردن f_t و ضرب درایه به درایه در ورودی C_{t-1} ، C_t اولیه محاسبه می شود.

$$C_t = C_{t-1} \odot f_t \quad \text{اثبات ۲-۲}$$

در اثبات ۲-۲، علامت $\odot$ به معنای ضرب نقطه ای دو بردار می باشد.

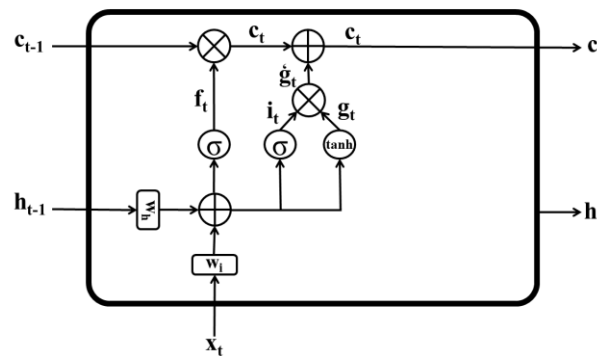
مرحله چهارم: در این مرحله به محاسبه خروجی g_t که اطلاعات استپ زمانی جاری را دارد، پرداخته می‌شود. این شبکه همانند گیت فراموشی شامل دو ورودی x_t و h_{t-1} است. بازهم این ورودی‌ها باید از دو لایه با وزن‌های w_{hg} و w_{ig} عبور کنند و بعد باهم جمع شوند. این ورودی از یک تابع تانژانت هایپربولیک باید عبور داده شود. خروجی g_t بین ۱- تا ۱ خواهد بود. برای اینکه ممکن است اثر تعدادی از درایه‌ها یا مولفه‌های موجود در C_t کاهش یابد. یعنی با مقادیر بین ۱- تا ۱ اثر بعضی مولفه‌ها را زیاد یا کم کرد (شکل ۹-۲). همچنین خروجی g_t از فرمول اثبات ۳-۲ محاسبه می‌شود:

$$g_t = \sigma (w_{hg} \cdot h_{t-1} + w_{ig} \cdot x_t + b_{hg} + b_{ig}) \quad \text{اثبات ۳-۲}$$



شکل ۹-۲ تشکیل اطلاعات استپ زمانی g_t

مرحله پنجم: گیت ورودی i_t ورود یک سری اطلاعات جدید به حافظه بلندمدت را بررسی می کند. به همین دلیل نامش را گیت ورودی گذاشته اند. مشابه گیت فراموشی، ممکن است مقادیر موجود در بردار i_t نزدیک به صفر باشد، بنابراین اثر g_t را کم می کند. بالعکس، ممکن است مقادیر بردار i_t نزدیک به ۱ باشد، در این صورت g_t در حافظه بلندمدت ذخیره می شود. ساختار این گیت دقیقاً مشابه با گیت فراموشی می باشد (شکل ۱۰-۲).



شکل ۱۰-۲ اضافه شدن گیت ورودی به ساختار داخلی شبکه LSTM

نحوه محاسبه گیت ورودی، خروجی نهایی g'_t و C_t نهایی:

$$i_t = \sigma(w_{hi}.h_{t-1} + w_{ii}.x_t + b_{hi} + b_{ii})$$

اثبات ۴-۲

$$\tilde{g}_t = g_t \odot i_t$$

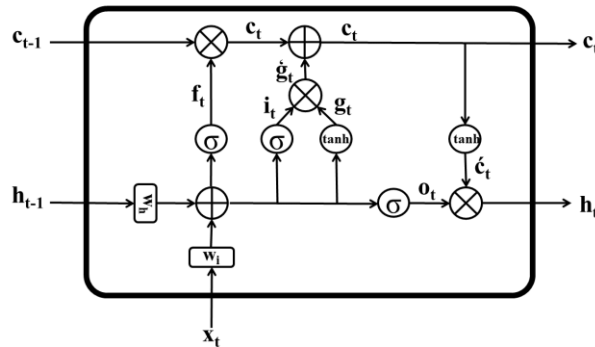
اثبات ۵-۲

$$c_t = c_t + \tilde{g}_t$$

اثبات ۶-۲

مرحله ششم: در این مرحله به ساختن خروجی h_t پرداخته می شود. در اولین قدم خروجی C_t را از یک تانژانت هایپربولیک عبور می کند، سپس به خروجی h_t متصل می شود. گیت خروجی o_t تعیین می کند چقدر از حافظه بلندمدت باید به خروجی منتقل شود. گیت خروجی، دو ورودی x_t و h_{t-1} را وارد دو لایه

می‌کند و این‌ها را باهم جمع می‌کند و در نهایت از تابع سیگموید عبور می‌دهد. در نهایت، خروجی تولید شده از گیت خروجی o_t در خروجی تابع سیگموید ضرب می‌شود تا مقداری که نیاز است به خروجی h_t منتقل شود (شکل ۱۱-۲).



شکل ۱۱-۲ ساخت خروجی h_t

نحوه محاسبه گیت خروجی o_t و خروجی h_t :

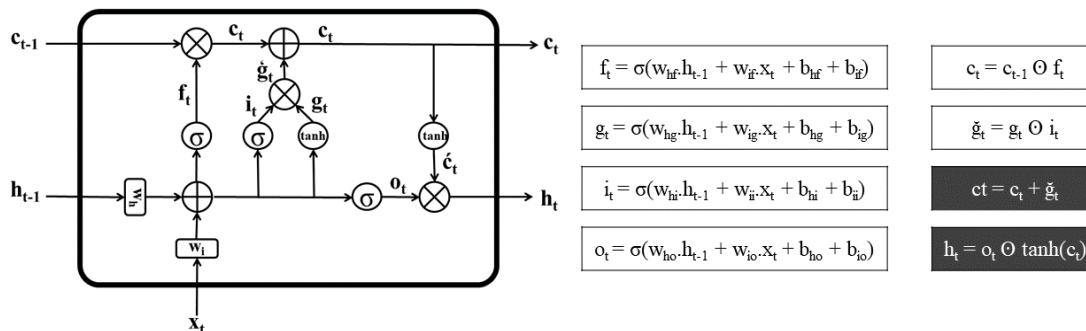
$$o_t = \sigma(w_{ho} \cdot h_{t-1} + w_{io} \cdot x_t + b_{ho} + b_{io})$$

اثبات ۷-۲

$$h_t = o_t \odot \tanh(c_t)$$

اثبات ۸-۲

شکل ۱۲-۲ جزئیات کامل و فرمول‌های مربوط به شبکه LSTM می‌باشد که به تفسیر و مرحله به مرحله توضیحات مربوط به آن ارائه شد.



شکل ۱۲-۲ نسخه کامل از شبکه عصبی LSTM همراه با تمامی فرمول‌های اثبات شده

۲ - ۳ پیشینه تحقیق

۲ - ۳ - ۱ روش پارامتریک (ARIMA)

برای تجزیه و تحلیل این داده‌ها و پیش بینی کوتاه مدت حجم تردد از مدل‌های مختلف مانند مدل خودهمبسته یک‌پارچه میانگین متحرک فصلی (SARIMA^۱) (گوین^۲، ۲۰۰۶)، مدل (ARIMA-GARCH^۳) (چن^۴، ۲۰۱۱)، مدل هموار سازی نمایی^۵ و مدل خودهمبسته یک‌پارچه میانگین متحرک (ARIMA) (گوش^۶، ۲۰۰۵)، مدل خودهمبسته یک‌پارچه میانگین متحرک (ARIMA) و مدل شبکه عصبی رگرسیون عمومی (GRNN^۷) (وو^۸، ۲۰۱۳) توسط چندین محقق اعمال می‌شود. بیشتر تکنیک‌های فوق برای داده‌های ترافیکی جمع‌آوری شده است و در تمام این بررسی‌ها به ارزش مدل خودهمبسته یک‌پارچه میانگین متحرک (ARIMA) برای داده‌های ترافیکی اتفاق نظر دارند.

گوش و همکاران، سه مدل سری زمانی مختلف، یعنی مدل پیاده‌روی تصادفی^۹، روش هموارسازی نمایی هولت وینترز^{۱۰} و مدل SARIMA برای مدل سازی جریان ترافیک استفاده کردند. حجم ترافیک را در فاصله زمانی پانزده دقیقه‌ای در تعدادی از مکان‌های مورد مطالعه ثبت کردند. از آنجا که رفتار سفر تعطیلات آخر هفته بسیار متفاوت از رفتار سفر در روزهای هفته است، اساساً مدل‌سازی را روی داده‌های روزهای هفته انجام دادند و داده‌های آخر هفته را حذف کردند. بعد از مقایسه سه روش مدل‌سازی مشاهده کردند که روش هموارسازی هولت وینترز و روش SARIMA بهترین مدل برای تقاطعات مورد مطالعه است در حالی که مدل SARIMA نیز کاملاً منطقی است (گوش^{۱۱}، ۲۰۰۵).

گوین در مقاله اش سعی به گسترش مدل پیش بینی SARIMA برای زمان سفر داشت. وی معتقد بود ابتکار عمل برای تحقیق در مورد مدل‌های پیچیده وجود دارد، اما مدل‌های ساده‌تر نیز می‌توانند سطح

^۱ Seasonal Autoregressive Integrated Moving Average Model (SARIMA)

^۲ Guin

^۳ Autoregressive Integrated Moving Average with Generalized Autoregressive Conditional Heteroscedasticity (ARIMA-GARCH)

^۴ Chen

^۵ Exponential smoothing

^۶ Ghosh

^۷ Generalized Regression Neural Network (GRNN)

^۸ Wu

^۹ random walk

^{۱۰} Holt-Winters' exponential smoothing

^{۱۱} Ghosh

بالایی از صحت را فراهم کنند. او نتیجه گرفت برای پیش بینی کوتاه مدت ، داده های حجم تردد چند سال اخیر کافی است ، اما برای پیش بینی های بلندمدت حجم تردد، داده های سنوات بیشتری مورد نیاز است و داده های محدود برای این منظور کافی نیست (گوین^۱، ۲۰۰۶).

چن و همکاران از مدل (ARIMA-GARCH) برای پیش بینی جریان ترافیک استفاده کردند. آن ها مدل ARIMA خطی را با مدل GARCH غیرخطی ترکیب کردند ، بنابراین هم میانگین و هم ناهمگونی شرطی را برای جریان ترافیک بررسی کردند و عملکرد مدل ترکیبی را با عملکرد مدل ARIMA استاندارد مقایسه کردند. نتایج نشان دادند که معرفی ناهمگونی نمی تواند پیشرفت مطلوبی در دقت پیش بینی به همراه آورد ، در بعضی موارد مدل GARCH حتی ممکن است عملکرد را بدتر کند. بنابراین برای پیش بینی جریان ترافیک معمولی ، مدل ARIMA استاندارد کافی است (چن^۲، ۲۰۱۱).

وو و همکاران ، دو مدل را آزمون کردند، یکی مدل ARIMA و دیگری مدل شبکه عصبی رگرسیون عمومی (GRNN) [۴]. هدف اصلی آن ها نشان دادن بهترین مدل پیش بینی ترافیک روزانه برای سیستم حمل و نقل هوشمند بود. در پایان آزمون دریافتند که مدل میانگین متحرک اتورگرسیون یک پارچه (ARIMA) بهتر از شبکه عصبی رگرسیون عمومی (GRNN) در مدل پیش بینی برای دوره های زمانی کوتاه کار می کند (وو^۳، ۲۰۱۳).

جهانگرد و همکاران (جهانگرد، ۱۳۹۵)، زمان سفر معابر درون شهری برای آینده را به کمک برآورد حجم تردد این معابر، پیش بینی کردند. داده های این تحقیق، برای چهار هفته متوالی گردآوری شده و برای مدلسازی از روش میانگین متحرک ساده استفاده شده است. به نظر می رسد در این تحقیق، ایستاسازی سری زمانی انجام نگرفته است، داده های تحقیق محدود می باشد و لذا نویسندگان برای مدل سازی سری زمانی به ساده ترین روش، یعنی روش میانگین متحرک ساده متوسل شده اند.

۲ - ۳ - ۲ مقایسه روش های پارامتریک و غیرپارامتریک

در چند دهه گذشته ، مدل های زیادی از تجزیه و تحلیل داده ها برای حل پیش بینی کوتاه مدت ترافیک ارائه شده است ، از جمله میانگین متحرک و هموارسازی (فرخی سعادت آبادی^۴، ۲۰۱۰) ، روش های آماری و رگرسیون (فی^۵، ۲۰۱۱)، روش های مبتنی بر نظریه جریان ترافیک (لی^۶، ۲۰۱۳) و تکنیک های

^۱ Guin

^۲ Chen

^۳ Wu

^۴ Farokhi Sadabadi

^۵ Fei

^۶ Li

یادگیری ماشین (وی^۱، ۲۰۱۲) این رویکردهای پیش‌بینی را می‌توان به دو دسته روش‌های پارامتریک^۲ و روش‌های غیرپارامتریک^۳ تقسیم کرد. در میان روش‌های پارامتریک، مدل ARIMA به طور گسترده به عنوان یک چارچوب پذیرفته شده برای ایجاد مدل پیش‌بینی ترافیک شناخته می‌شود. بسیاری از پژوهش‌های مربوط به ARIMA در دهه‌های گذشته انجام شده‌است. در اوایل دهه ۱۹۸۰، لوین و همکاران، برای پیش‌بینی جریان ترافیک آزادراه از تجزیه و تحلیل سری‌زمانی Box-Jenkins استفاده کردند و دریافتند که مدل $ARIMA(0, 1, 1)$ از نظر آماری بیشترین اهمیت را دارد (لوین^۴، ۱۹۸۰). در همان دوره، احمد و همکاران، مدل ARIMA را برای پیش‌بینی جریان ترافیک در جاده‌های شریانی شهری استفاده کردند (احمد^۵، ۱۹۷۹). ARIMA که از لحاظ نظری به خوبی تعریف شده و عملاً موثر است، به تدریج یک معیار برای مقایسه مدل‌های پیش‌بینی جدید توسعه یافت. وقتی ترافیک تغییرات منظم را نشان می‌دهد، روش‌های پارامتریک می‌توانند عملکرد خوبی را به دست آورند، اما زمانی که ترافیک تغییرات نامنظم را نشان می‌دهد، خطای پیش‌بینی آشکار است. برای حل این مشکل، محققان به روش‌های غیرپارامتریک در زمینه پیش‌بینی جریان ترافیک، مانند رگرسیون غیرپارامتریک، شبکه عصبی، ماشین بردار پشتیبان (SVM)^۶، فیلتر کالمن^۷ توجه زیادی داشتند (ژائو^۸، ۲۰۱۷).

مدل‌های پیش‌بینی موجود به سه دسته طبقه‌بندی می‌شوند: روش‌های پارامتریک، روش‌های غیرپارامتری و روش‌های ترکیبی. مدل‌های روش‌های پارامتری شامل مدل خودهمبسته یک‌پارچه میانگین متحرک (ARIMA) (وان در ورت^۹، ۱۹۹۶)، مدل خودهمبسته یک‌پارچه میانگین متحرک فصلی (SARIMA) (ویلیامز^{۱۰}، ۲۰۰۳) و فیلتر کالمن^{۱۱} (گوو^{۱۲}، ۲۰۱۴) است. روش‌های پارامتری به طور گسترده‌ای در پیش‌بینی جریان ترافیک استفاده می‌شوند، اما این روش‌ها در شرایط مختلف به داده‌های ترافیک وابسته هستند. مدل‌های روش‌های غیرپارامتری شامل مدل شبکه‌های عصبی مصنوعی (ANN) (چن^{۱۳}، ۲۰۱۹)، ضریب K نزدیکترین همسایه^{۱۴} (KNN) (برناس^{۱۵}، ۲۰۱۵)، رگرسیون بردار

^۱ Wei

^۲ parametric

^۳ nonparametric

^۴ Levin

^۵ Ahmed

^۶ Support Vector Machines

^۷ Kalman Filter

^۸ Zhao

^۹ Van Der Voort

^{۱۰} Williams

^{۱۱} Kalman filter

^{۱۲} Guo

^{۱۳} Chen

^{۱۴} K-Nearest Neighbor

^{۱۵} Bernas

پشتیبان^۱ (SVR) (کاسترو نتو^۲، ۲۰۰۹) و مدل بیزی^۳ (وانگ^۴، ۲۰۱۴) است. در مقایسه با روش‌های پارامتریک، روش‌های غیرپارامتری در پیش‌بینی عملکرد موثرتری دارند. با این وجود، روش‌های غیرپارامتری به مقدار زیادی داده و فرایند آموزش نیاز دارد (لوو^۵، ۲۰۱۹).

۲ - ۳ - یادگیری عمیق

تیان و پان^۶ با استفاده از شبکه LSTM یک لایه پنهان با ۵ تا ۴۰ واحد را پیاده‌سازی کرده و این مدل را با استفاده از سیستم اندازه‌گیری عملکرد کالیفرنیا روی داده‌های ۶ آزادراه در مدت ۲۴۹ روز کاری در سال ۲۰۱۴ در فواصل ۵ دقیقه‌ای ارزیابی کردند. آن‌ها نتایج LSTM را با مدل‌های پیاده‌روی تصادفی، ماشین بردار پشتیبان شعاعی، شبکه عصبی تک‌لایه پیشخور و خود رمزگذار انباشته (SAE)^۷ مقایسه می‌کنند. نتایج این مقایسه نشان داد که LSTM دارای درصد میانگین مطلق خطای^۸ و مجذور میانگین مربعات خطای^۹ کمتر برای همه فواصل آزمایش را دارد (تیان، ۲۰۱۵).

ژائو و همکاران^{۱۰} داده‌های حجم تردد خودرو را در یکی از کمربندی‌های مهم شهر پکن چین توسط دوربین‌های دارای سیم پیچ القایی و رادار سرعت که داده‌های از دست رفته کتری دارند برداشت نمودند. آن‌ها از داده‌های ماه اول سال ۲۰۱۵ میلادی استفاده کردند. ۵ ماه اول را به عنوان داده‌های آموزش و یک ماه پایانی را به عنوان داده‌های تست کنار گذاشتند. آن‌ها با توجه به مقایسه روش LSTM با دیگر متدولوژی‌های پیشرفته، نتیجه گرفتند که مدل شبکه LSTM پیشنهادی برای پیش‌بینی حجم ترافیک قوی است (ژائو، ۲۰۱۷).

جیا و همکاران^{۱۱} با استفاده از روش‌های یادگیری عمیق با معماری‌های مدل پیچیده‌تر و قابلیت‌های داده‌کاوی مؤثرتر، شبکه باور عمیق^{۱۲} (DBN) و حافظه کوتاه‌مدت (LSTM) را برای پیش‌بینی جریان ترافیک شهری با توجه به تأثیر بارندگی معرفی می‌کند. با توجه به تأثیر بارندگی، توافق عمومی وجود دارد که به طور قابل توجهی بر ویژگی‌های جریان ترافیک تأثیر می‌گذارد و منجر به ازدحام و تصادف می‌شود. بدون درک کلی از تأثیر آب و هوا بر جریان ترافیک، مدیران مدیریت ترافیک نمی‌توانند عوامل مرتبط را در روش‌های کارآمد مرتبط برای بهبود کارایی و ایمنی ترافیک در نظر بگیرند. هدف آن‌ها

^۱ support vector regression

^۲ Castro-Neto

^۳ Bayesian model

^۴ Wang

^۵ Luo

^۶ Tian and Pan

^۷ Stacked Autoencoder

^۸ Mean Absolute Percentage Error (MAPE)

^۹ Root Mean Squared Errors (RMSE)

^{۱۰} Zheng Zhao

^{۱۱} Yuhan Jia et al.

^{۱۲} Deep Belief Network

مقایسه تفاوت بین شبکه عصبی مکرر پیشرفته و یادگیری عمیق برای پیش‌بینی جریان ترافیک سری زمانی است. آن دریافتند که DBN و LSTM می‌توانند به طور موثر ویژگی‌های جریان ترافیک را تحت شرایط مختلف بارندگی کشف کنند، و مدل‌های جدید می‌توانند عملکرد بهتری از BPNN^۱ و ARIMA داشته باشند و همچنین نسبت به مدل‌های اصلی یادگیری عمیق بدون بارندگی بهبودهایی را به همراه داشته باشند. علاوه بر این، یادگیری عمیق LSTM می‌تواند برای توصیف ویژگی‌های سری زمانی داده‌های ترافیکی از شبکه عصبی DBN بهتر عمل کند. به طور خلاصه، نتایج اهمیت ترکیب تاثیر آب و هوا و پیشرفت‌های قابل توجهی در دقت پیش‌بینی را دریافتند. این کار نشان می‌دهد که یادگیری عمیق با ورودی‌های داده چند منبعی امیدوارکننده است (جیا، ۲۰۱۷).

روش‌های ترکیبی عمدتاً ترکیب روش پارامتریک با روش غیرپارامتریک هستند. اگرچه دقت پیش‌بینی روش‌های غیرپارامتری و روش‌های ترکیبی نسبت به روش‌های پارامتریک برتری دارند، اما همه این روش‌ها عمدتاً داده‌های مربوط به ایستگاه پیش‌بینی را در نظر می‌گیرند، که نمی‌تواند ویژگی‌های زمانی-مکانی داده‌های جریان ترافیک را به طور کامل نشان دهد. به تازگی، یادگیری عمیق، یک روش یادگیری ماشینی در حال ظهور، توجه زیادی را از جانب دانشگاه و صنعت جلب کرده است. پیش‌بینی جریان ترافیک بر اساس روش‌های یادگیری عمیق به یک روند جدید تبدیل شده است. با این حال، شبکه‌های RNN با مشکل وابستگی بلندمدت مواجه هستند. برای حل این مشکل، شبکه‌های LSTM پیشنهاد شدند. از آنجا که LSTM می‌تواند به طور خودکار وقعه‌های زمانی مطلوب را محاسبه کرده و ویژگی‌های سری‌های زمانی را با بازه زمانی طولانی‌تر ثبت کند، می‌توان با مدل LSTM در پیش‌بینی جریان ترافیک عملکرد بهتری را به دست آورد. ما و همکاران^۲ (ما، ۲۰۱۹) شائو و سونگ^۳ (شائو، ۲۰۱۶) از LSTM برای یادگیری تجربیات بیشتر در داده‌های جریان ترافیک غیرخطی برای جلب وابستگی زمانی طولانی مدت به دنباله‌های ترافیک استفاده کردند (لوو، ۲۰۱۹).

خان و همکاران^۴ برای ۸۳ راه بین ایالتی، ۵۹ راه اصلی، ۱۵ مسیر مواصلاتی و ۱۵ راه دسترسی محلی از روش‌های RNN^۵، GRU^۶ و LSTM برای پیش‌بینی حجم تردد ساعتی برای ۳۶۵ روز سال بعد است تا بتوان برآورد دقیقی از AADT^۷ سال آینده را از داده‌های سال‌های قبل به دست آورد. هدف خان و همکاران توسعه مدلی را برای پیش‌بینی AADT می‌باشد که وابستگی‌های بلندمدت در مجموعه داده‌های ترافیکی را نشان می‌دهد. این مطالعه همچنین مدل LSTM را به عنوان بهترین مدل برای پیش‌بینی دقیق AADT در حالی که تغییرات فصلی بلندمدت در سری‌های زمانی را ثبت می‌کند،

^۱ Back-Propagation Neural Network

^۲ Ma et al.

^۳ Shao and Soong

^۴ Zaid Khan et al.

^۵ Recurrent Neural Network (RNN)

^۶ Gated Recurrent Unit (GRU)

^۷ Annual Average Daily Traffic (AADT)

شناسایی کند. به طور کلی، مدل LSTM با میانگین $RMSE = 274$ و $MAPE = 18.91\%$ برای پیش‌بینی AADT و میانگین $RMSE = 824$ و $MAPE = 2.10\%$ برای پیش‌بینی حجم ترافیک ساعتی، بهترین مدل در تمام مدل‌ها برای پیش‌بینی AADT در نظر گرفته است. این مدل می‌تواند تغییرات طولانی‌مدت در مجموعه داده را در حالی که دقت بالای AADT و پیش‌بینی حجم ترافیک ساعتی را حفظ می‌کند، ثبت کند. این مدل همچنین قادر به پیش‌بینی چندساله با کاهش جزئی در دقت است (خان، ۲۰۱۹).

لو و همکاران^۱ یک روش پیش‌بینی جریان ترافیک مکانی-زمانی را با KNN^۲ و LSTM پیشنهاد کردند. KNN برای انتخاب ایستگاه‌های همسایه که بیشترین ارتباط فضایی-زمانی را با ایستگاه منتخب نشان می‌دهند، استفاده کردند. یک شبکه LSTM برای پیش‌بینی جریان ترافیک به ترتیب در ایستگاه-های منتخب استفاده کردند. LSTM قادر است از وابستگی بلندمدت در داده‌های جریان ترافیک استفاده کرده و ویژگی‌های پنهان در جریان ترافیک را کشف کند، که عملکرد پیش‌بینی بهتری را ارائه می‌دهد. آن‌ها عملکرد مدل خود را با داده‌های ترافیکی واقعی ارائه شده توسط TDRL^۳ ارزیابی کردند و با مدل ARIMA^۴، SVR^۵، WNN^۶، DBN^۷ و LSTM مقایسه کردند. نتایج نشان دادند که مدل پیشنهادی نسبت به سایر روش‌ها برتری دارد. از آنجا که داده‌های جریان ترافیک تحت تأثیر آب و هوا، حوادث و عوامل دیگر قرار می‌گیرد، تأثیر این عوامل بر داده‌های جریان ترافیک بیشتر مورد مطالعه قرار می‌گیرد تا دقت پیش‌بینی بهبود یابد (لو، ۲۰۱۹).

دوگان^۸ مدل‌های برآورد جریان ترافیک با استفاده از روش‌های هوش مصنوعی و شبکه‌های عمیق توسعه داد و اندازه مجموعه‌های آموزشی (تعداد داده‌ها) مورد بررسی قرار داد. او در این مطالعه روش شبکه‌های عصبی اتورگرسیو غیرخطی^۹ (NAR) و یادگیری عمیق (LSTM) را مورد بررسی قرار داد. او دریافت اندازه مجموعه آموزشی بزرگتر همیشه به معنای عملکرد بهتر برای LSTM و NAR نیست. مقدار زیاد مجموعه آموزشی باعث افزایش کارایی نمی‌شود. به همین دلیل، اندازه مجموعه آموزشی بهینه در روش‌های یادگیری عمیق باید تعیین شود. همچنین دریافت که LSTM کمتر از مدل NAR تحت تأثیر تغییرات آنی جریان ترافیک قرار می‌گیرد. از نظر آماری نیز، زمانی که اندازه مجموعه آموزشی بزرگتر از ۴۸۰ باشد، روش LSTM بهتر از NAR عمل می‌کند. بنابراین، مدل LSTM نتایج پایداری را نسبت به NAR برای پیش‌بینی ترافیک کوتاه‌مدت تولید می‌کند (دوگان، ۲۰۲۰).

^۱ Xianglong Luo et al.

^۲ K - Nearest Neighbor

^۳ Research Data Lab

^۴ Support Vector Regression

^۵ Wavelet Neural Network

^۶ Deep Belief Networks

^۷ Erdem DOGAN

^۸ Nonlinear Autoregressive Neural Networks

۲ - ۳ - ۴ جمع بندی

مدل ARIMA اولین بار در سال ۱۹۷۹ توسط احمد و همکاران برای پیش‌بینی جریان ترافیک در جاده‌های شریانی شهری مورد استفاده قرار گرفت و دریافتند که این روش از لحاظ نظری به خوبی تعریف شده است و عملاً در پیش‌بینی جریان ترافیک موثر است. این مدل به تدریج در سال‌های بعدی مورد توجه قرار گرفت و به عنوان یک معیار مقایسه مدل‌های پیش‌بینی جدید توسعه یافت.

در دهه‌های اخیر با توجه مهم‌ترین نقطه ضعف مدل ARIMA و زیرشاخه‌های آن (فرض خطی بودن در سری‌زمانی) استفاده از این مدل را در بسیاری از مسائل عملی محدود کرد. پژوهشگران در حوزه پیش‌بینی برای حل این مشکل سعی به استفاده از روش‌های جایگزین کردند. اما از آنجایی که این روش سال‌ها به عنوان معیار مقایسه مدل‌های پیش‌بینی بوده است نمی‌توان این روش را به سادگی فراموش کرد. پس برای روش‌های جدید مدل ARIMA یک مدل مقایسه‌ای قرار گرفت.

در چند دهه گذشته روش‌های غیرپارامتریک مورد توجه قرار گرفت و این روش‌ها نسبت به روش‌های پارامتریک در پیش‌بینی عملکرد موثرتری داشتند و همچنین مشکلات اصلی این روش‌ها در سال اولیه نیاز به مقداری داده و نیاز به فرایند آموزش بود. در سال‌های اخیر تقریباً این مشکلات برطرف شده است. اما پژوهشگران دریافتند که روش‌های غیرپارامتریک دارای معماری سطحی می‌باشند و در استخراج همبستگی داده‌ها عملکرد ضعیفی دارند.

با ظهور یادگیری عمیق این روش مورد توجه زیادی از سمت پژوهشگران و صنعتگران در امر پیش‌بینی قرار گرفت. روش‌های یادگیری عمیق در سال‌های اولیه خود قرار دارند پس برای اثبات برتری روش یادگیری عمیق نسبت به روش‌های پارامتریک و غیرپارامتریک مطالعات گسترده‌ای در این زمینه انجام شود.

در این پژوهش سعی شده است از روش‌های پارامتریک، غیرپارامتریک و یادگیری عمیق برای پیش‌بینی جریان ترافیک استفاده شود. همچنین برای هر روش از مدل‌های معتبر، پرکاربرد و نوین استفاده شود. از آنجایی که داده‌های پیش‌بینی از نوع داده‌های سری‌زمانی می‌باشد و مولفه زمان از اهمیت زیادی برخوردار است پس بهترین روش یادگیری عمیق در این زمینه یعنی LSTM مورد بررسی قرار گرفته است.

فصل ۳: روش تحقیق

۳ - ۱ مقدمه

دو نوع از مهم‌ترین و قدرتمندترین روش‌های مدلسازی و پیش‌بینی سری‌های زمانی روش‌های مبتنی بر شبکه‌های عصبی بازگشتی و روش‌های مدلسازی ARIMA است. از آنجایی که اعتبار یک پژوهش به روش‌های بکار گرفته‌شده در آن سنجیده می‌شود. در پژوهش حاضر سعی گردیده تا از روش‌های مناسب و شناخته شده در پیش‌بینی استفاده گردد. در پژوهش حاضر برای پیش‌بینی داده‌های سری زمانی حجم تردد جاده‌ای از روش‌های پارامتریک، غیرپارامتریک و یادگیری عمیق استفاده می‌گردد. برای روش پارامتریک از روش باکس جنکینز (ARIMA فصلی) که شناخته شده و کارآمدترین روش پیش‌بینی در این روش می‌باشد، استفاده می‌گردد و برای روش غیرپارامتریک از روش شبکه عصبی مصنوعی و روش یادگیری عمیق (LSTM) که از بروزترین و مناسب‌ترین روش پیش‌بینی می‌باشد، استفاده می‌گردد.

هدف از مدل‌های پیش‌بینی در این پژوهش رسیدن به یک روش پیش‌بینی با کمترین درصد خطا، سریع‌ترین روش و مطمئن‌ترین روش برای طراحی و برنامه ریزی و کنترل ترافیک در بخش‌های مختلف جاده‌های کشور می‌باشد.

۳ - ۲ روش ARIMA و نرم افزار SPSS

روش ARIMA به سادگی مجموعه داده‌های استاندارد را در داده‌های سری زمانی فراهم می‌کند و بدین ترتیب یک راه‌کار ساده و در عین حال قدرتمند به منظور پیش‌بینی تخصصی سری‌های زمانی برای داده‌های خطی ارائه می‌دهد.

هر یک از این مولفه‌های تعریف شده در این روش به طور مشخص در مدل به عنوان یک پارامتر در نظر گرفته شده‌اند. همچنین این مدل برای هر یک از پارامترها از علامت استاندارد $ARIMA(p,d,q)(P,D,Q,S)$ استفاده می‌کند که در آن پارامترها با مقادیر صحیح جایگزین می‌شوند تا در مدل ARIMA ارائه شود.

پارامترهای مدل ARIMA به شرح زیر تعریف می‌شوند:

- P, p : تعداد مشاهدات موجود در مدل که ترتیب تأخیر نیز نامیده می‌شود.
- D, d : تعداد دفعاتی که مشاهدات خام از یکدیگر متمایز می‌شوند، درجه تمایز نیز نامیده می‌شود.
- Q, q : اندازه میانگین متحرک است.

۳-۲-۱ مراحل انجام پیش بینی در نرم افزار SPSS به روش ARIMA

به طور کلی در این روش پس از تعیین مرتبه تفاضلی کردن و تعیین مرتبه هر یک از فرایندهای AR و MA پارامترهای مدل مشخص می گردد. بررسی مناسبت مدل با تجزیه و تحلیل باقیمانده های مدل برازش داده شده صورت می گیرد. چنانچه مدل درست تشخیص داده شده باشد، باقیمانده ها باید دارای خواص متغیرهای تصادفی نرمال مستقل با میانگین صفر و واریانس ثابت باشند.

۳-۲-۱-۱ ورود داده ها و ساختن سری مانی

اولین مرحله در نرم افزار SPSS وارد کردن داده های سری زمانی مورد نظر و تعریف متغیرها در نرم افزار مانند نام متغیر، تعیین برچسب متغیر و... می باشد. در قدم بعدی برای داده های مورد نظر که وابسته به زمان هستند، اقدام به ساختن سری زمانی در نرم افزار می شود. در این مرحله براساس زمان جمع آوری داده ها (روز- هفته) را برای نرم افزار تعریف می نماییم که این داده ها از چه روزی و هفته ای آغاز می شوند.

۳-۲-۱-۲ تشخیص ایستایی داده ها

ایستایی مبحث بسیار مهمی در مدلسازی سری های زمانی می باشد. زیرا بسیاری از مدل های احتمالی سری های زمانی بر مبنای ایستایی سری استوار هستند. بنابراین اولین قدم در پیش بینی سری زمانی تشخیص ایستا بودن سری زمانی است. یک سری زمانی را سری ایستا یا مانا می نامیم هرگاه مشخصه های آماری آن مانند میانگین و واریانس در طول زمان ثابت بمانند.

۳-۲-۱-۳ هموارسازی سری نایستا

در این مرحله با استفاده از نمودار سری زمانی داده ها به شناسایی روند، نایستایی، فصلی بودن و شناسایی داده های پرت پرداخته می شود. اولین اقدام بررسی ایستایی در میانگین که با استفاده از نمودارهای تابع خودهمبستگی (ACF) و تابع خودهمبستگی جزئی (PACF) انجام می گیرد. نمودار ACF نشان می دهد که داده باقیمانده تحت تاثیر چند داده قبل قرار گرفته و مرتبه اتورگرسیون را مشخص می کند. همچنین نمودار PACF نشان می دهد که مجموع چند داده باقیمانده صفر می شود و مرتبه میانگین متحرک را مشخص می کند.

۳-۲-۱ تشخیص مدل (برآورد ضرایب)

در این مرحله با استفاده از نمودارهای سری زمانی، ACF و PACF اقدام به برآورد ضرایب p ، d و q برای مولفه‌های غیر فصلی و P ، D و Q برای مولفه‌های فصلی می‌نماییم.

۳-۲-۱-۵ برازش مدل

در این مرحله برازش مدل به معنی برآورد پارامترهای مجهول می‌باشد. نرم افزار SPSS با استفاده از یک روش تکراری، پارامترها را برآورد می‌نماید.

۳-۲-۱-۶ بررسی مناسب بودن مدل

برای مناسب بودن مدل می‌توان از روش تجزیه و تحلیل داده‌ها استفاده نمود. برای این منظور می‌توان مقادیر باقیمانده هریک از داده‌ها را به دست آورد و به بررسی نمودارها و آزمون‌های مربوط به آن‌ها پرداخت.

۳-۲-۱-۷ پیش بینی نهایی

در صورت مناسب بودن مدل برای برآورد اولیه ضرایب، این مدل تایید می‌شود و می‌توانیم آن را مبنای پیش‌بینی رفتار آینده قرار دهیم.

۳ - ۳ روش شبکه عصبی و یادگیری عمیق LSTM

در پژوهش حاضر، مدل شبکه عصبی مصنوعی و یادگیری عمیق LSTM به روش‌های مشابه کدنویسی شده‌اند، با این تفاوت که هر مدل به روش خاص خود آموزش می‌بیند. یادگیری عمیق را عمدتاً می‌توان همان شبکه عصبی عمیق بنامیم. تفاوت یادگیری عمیق و شبکه عصبی در آن است که یادگیری عمیق محدوده‌ای گسترده‌تر از شبکه عصبی را دارد و الگوریتم‌های یادگیری تقویتی^۱ را نیز دربرمی‌گیرد. تفاوت دیگر شبکه عصبی و یادگیری عمیق در تعداد لایه‌های درونی آنهاست. شبکه عصبی حداکثر از سه لایه تشکیل شده‌است در حالی که در شبکه یادگیری عمیق داده‌ها باید از یک شبکه با چندین لایه عبور کنند. همچنین در یادگیری عمیق هر لایه بر روی ویژگی خاصی از داده کار می‌کند و شبکه قادر به تشخیص ویژگی‌های پیچیده‌تری می‌باشد.

^۱ Reinforcement learning

۳-۳-۱ برنامه‌نویسی شبکه عصبی مصنوعی و یادگیری عمیق

LSTM در نرم‌افزار MATLAB

در پژوهش حاضر برای پیش‌بینی داده‌های سری زمانی موجود به روش شبکه عصبی مصنوعی و یادگیری عمیق LSTM از نرم‌افزار MATLAB جهت کد نویسی استفاده می‌گردد. مراحل مختلف کدنویسی و مدل‌سازی به منظور پیش‌بینی حجم تردد روزانه جاده‌ای شرح داده می‌شود.

۳-۳-۱-۱ فراخوانی داده‌های حجم تردد روزانه

در ابتدای برنامه نویسی، داده‌های مورد نظر که به صورت اعداد اسکالر که نشان‌دهنده تعداد خودرو عبوری از محور مورد نظر می‌باشد، از قدیمی‌ترین روز (اولین روز موجود) به صورت پشت‌سرهم در نرم‌افزار اکسل قرار می‌گیرند و به صورت یک فایل اکسل وارد برنامه می‌گردد.

بعد از فراخوانی داده‌های سری‌زمانی، به وابستگی داده‌ها به روزهایی که به آن وابسته است و از آن تاثیر می‌پذیرد، پرداخته می‌شود. در پژوهش حاضر فرض شده‌است که داده‌های هر روز به ۰ (امروز)، ۱، ۲، ۳، ۴، ۵، ۶، ۷، ۱۴، ۳۰، ۳۱، ۳۵۳، ۳۵۴، ۳۵۵، ۳۶۴ و ۳۶۵ روز قبل وابستگی دارد و از آن تاثیر می‌پذیرد. با جداسازی داده‌ها بدین صورت به جای یک ماتریس $1 \times n$ (فایل فراخوانی شده)، یک ماتریس $16 \times n$ خواهیم داشت که مراحل بعدی برنامه نویسی بر روی این ماتریس صورت می‌گیرد.

۳-۳-۱-۲ نرمال‌سازی داده‌ها^۱

نرمال‌سازی داده‌ها عبارت است از روشی که داده‌ها را در زمانی که در یک دامنه نیستند را در دامنه مشابه قرار می‌دهد. روش‌های مختلفی جهت نرمال‌سازی داده‌ها وجود دارند. در این پژوهش یکی از معروف‌ترین این روش‌ها را به MinMaxNormalization معروف است استفاده می‌شود. طبق فرمول کلی این رابطه داده‌ها عددی بین ۱- و ۱ قرار می‌گیرند.

در این مرحله به نرمال‌سازی داده‌ها پرداخته می‌شود. عمل نرمال‌سازی در کدنویسی حاضر به صورت دستی صورت گرفته‌است که براساس مقادیر مینیمم و ماکزیمم در هر ستون می‌باشد. تابع نرمال به صورت رابطه ۳-۱ می‌باشد.

$$X_n = [(X - \text{Minx}) / (\text{Maxx} - \text{Minx}) * 2] - 1; \quad \text{رابطه ۳-۱}$$

در رابطه ۳-۲ X مقدار نرمال شده و عددی بین ۱- تا ۱، X_n برابر با حجم تردد روزانه، Maxx مقدار ماکزیمم حجم تردد در هر محور، Minx مقدار مینیمم حجم تردد در هر محور می‌باشد.

^۱ data Nnormalization

۳-۱-۳-۳ جداسازی داده‌های آموزش و آزمایش

در این مرحله داده‌های نرمال شده را به دو قسمت آموزش و آزمایش تقسیم می‌شود. بدین صورت که داده‌های از روز اول تا یکسال پایانی به عنوان داده‌های آموزش و داده‌های یکسال پایانی (۳۶۵ روز) به عنوان داده‌های آزمایش در نظر گرفته می‌شود. مدل با داده‌های آموزش، آموزش می‌بیند و سپس با داده‌های آزمایش اعتبارسنجی می‌شوند.

۳-۱-۳-۴ تشکیل ساختار شبکه

این مرحله تمایز میان شبکه عصبی و یادگیری عمیق LSTM می‌باشد. شبکه عصبی با دولایه که در لایه اول ۱۰ نورون و در لایه دوم ۵ نورون می‌باشد، تشکیل می‌شود. همچنین برای LSTM تعداد واحدهای پنهان ۲۰۰، بیشترین مقدار تکرار ۵۰، نرخ یادگیری اولیه ۰,۰۰۵ و ضریب کاهشی نرخ یادگیری ۰,۲ می‌باشد. در تشکیل ساختار شبکه‌های موجود برای داده‌های مورد نظر بهینه‌ترین حالت در نظر گرفته شده‌است.

۳-۱-۳-۵ آموزش و ارزیابی شبکه

در این مرحله شبکه با استفاده از ساختار شبکه، ماتریس ورودی و خروجی (x و y) داده‌های آموزش، آموزش می‌بیند و ماتریس ورودی و خروجی داده‌های آزمایش شبکه (داده‌های پیش‌بینی) را حاصل می‌شود. از آنجایی که داده‌های مورد نظر به صورت نرمال در شبکه وجود دارند می‌بایستی از عکس تابع نرمال (رابطه ۳-۲) استفاده شود تا داده‌های آزمایش را بتوان با داده‌های اصلی مقایسه و اعتبارسنجی نمود.

$$X = [(X_n + 1) * (Maxx - Minx) * 0,5] + Minx; \quad \text{رابطه ۳-۲}$$

۳-۱-۳-۶ اطلاعات خروجی نرم افزار

در این مرحله، پس از آموزش شبکه اطلاعات خروجی نرم‌افزار مانند اشکال، نمودارها و جداول داده‌های پیش‌بینی استخراج می‌شود و مورد ارزیابی قرار می‌گیرند. همچنین نمودارهای مقایسه‌ای از داده‌های اصلی و پیش‌بینی شده با استفاده از نرم‌افزار Excel ترسیم و در فصل ۴ ارائه می‌شود.

۳-۱-۷ اعتبار سنجی

در این مرحله به اعتبار سنجی و محاسبه میزان خطا داده های آزمایش پرداخته می شود. برای این منظور از فرمول میانگین مربعات خطا (MSE^1) (رابطه ۳-۳)، مجذور میانگین مربعات خطا ($RMSE^2$) (رابطه ۴-۳) و میانگین درصد خطای مطلق ($MAPE^3$) (رابطه ۵-۳) استفاده گردیده است. فرمول این فرمول ها در روابط زیر آمده است:

$$MSE = \frac{1}{N} \sum_{i=1}^N (\hat{y}(i) - y(i))^2 \quad \text{رابطه ۳-۳}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}(i) - y(i))^2} \quad \text{رابطه ۴-۳}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{\hat{y} - y}{y} \right| * 100 \quad \text{رابطه ۵-۳}$$

در روابط فوق N تعداد کل نمونه، $\hat{y}$ حجم تردد پیش بینی شده، y مقدار حجم تردد اصلی می باشد. همچنین با استفاده از داده های حجم متوسط روزانه در سال (AADT) برای سال افق پیش بینی (۱۳۹۸) و با استفاده از درصد دقت پیش بینی 4 (رابطه ۶-۳) می توان سه مدل ARIMA، شبکه عصبی مصنوعی و یادگیری عمیق LSTM را بایکدیگر مقایسه نمود.

$$Accuracy = (100 - MAPE_{AADT}) \% \quad \text{رابطه ۶-۳}$$

۳ - ۴ نتیجه گیری

روش های ارائه شده در این بخش از مهم ترین و قدرتمندترین روش های مدلسازی و پیش بینی سری های زمانی می باشد. روش LSTM، یک روش نوین، با دقت بالا و کارآمد در امر پیش بینی سری زمانی می باشد که در چند سال گذشته در این زمینه از محبوبیت خاصی برخوردار بوده است.

^۱ Mean Square Error (MSE)

^۲ Root Mean Square Error (RMSE)

^۳ Mean Absolute Percentage Error (MAPE)

^۴ Accuracy

فصل ۴ : مدلسازی و تحلیل نتایج

۴ - ۱ مقدمه

در پژوهش حاضر از اطلاعات سازمان راهداری و حمل و نقل جاده‌ای کشور استفاده گردیده‌است. اطلاعات ترددشماری از سال ۱۳۷۴ به صورت شمارش بصری، از سال ۱۳۸۱ به صورت مکانیزه از سیستم حلقه القایی دفنی در جاده‌ها و از سال ۱۳۸۹ اطلاعات به صورت برخط در سامانه ترددشماری برداشت و گزارش شده است. از سال ۱۳۸۹ در پایان هر ماه اطلاعات آفلاین هر محور ترددشماری به صورت روزانه و ساعتی بر روی سایت قرار می‌گیرد. براساس قرداد، اطلاعات ترددشماری باید دقتی در حد ۹۵٪ داشته باشد اما با توجه به مشکلات احتمالی به طور متوسط دقتی حدود ۹۰٪ از کل اطلاعات برآورد شده مورد انتظار می‌باشد.

در پژوهش حاضر از اطلاعات مربوط به استان خراسان رضوی استفاده گردیده‌است. با توجه به اهمیت مسیرهای موجود در این استان، ۶ مسیر رفت و برگشت دارای اهمیت و تردد بالاتر در نظر گرفته می‌شود این مسیرها عبارتند از:

- مشهد - سرخس
- مشهد - چناران
- مشهد - فریمان
- مشهد - ملک آباد
- مشهد - کلات
- مشهد - سد کارده

بنابراین در این فصل به معرفی مسیرهای مورد ارزیابی جهت چیش بینی حجم تردد و مدلسازی، نحوه گردآوری داده‌ها، مدلسازی و پیش‌بینی به هریک از روش‌ها، مقایسه و نتیجه‌گیری پرداخته می‌شود.

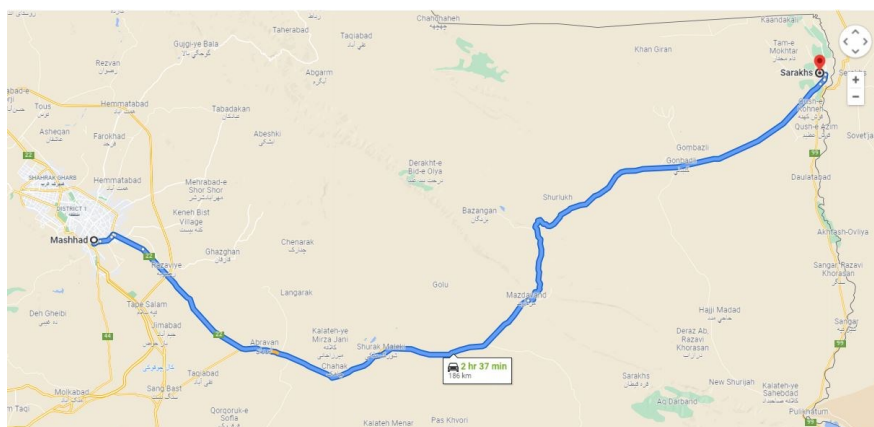
۴ - ۲ معرفی مسیرهای مورد ارزیابی

خراسان رضوی و مشهد به عنوان مقصد اصلی سفرهای جاده‌ای، سالانه دست کم ۲۰ میلیون مسافر، می‌تواند شاخصی برای ارزیابی اقدامات حمل و نقل جاده‌های کشور در نظر گرفته شود. این استان با ۱۱۰ کیلومتر آزادراه در رتبه نهم کشور، ۷۲ کیلومتر بزرگراه در رتبه پنجم، یک هزار و ۲۸۹ کیلومتر راه اصلی در رتبه ششم، سه هزار و ۹۵۰ کیلومتر راه فرعی در رتبه چهارم در کشور دارا می‌باشد. همچنین این استان با بیست هزار و ۵۷۴ کیلومتر طول کل راه‌های برون‌شهری بعد از استان فارس در

رتبه دوم قرار دارد. از این جهات است که پیش بینی و برنامه‌ریزی جریان ترافیک و حجم تردد در مسیرهای این استان حائز اهمیت می‌باشد.

مسیر مشهد – سرخس

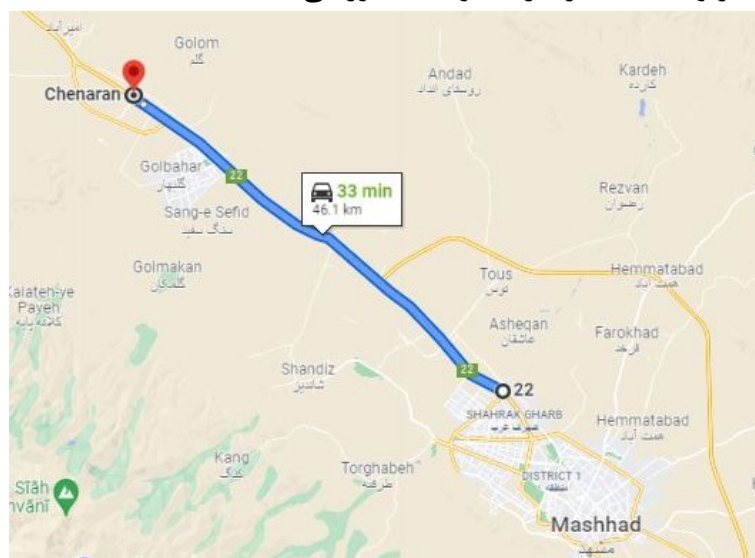
این جاده یک محور ترانزیتی محسوب می‌شود و مشهد از طریق این جاده به سرخس و شهر سرخس را به شهر سرخس ترکمنستان ارتباط می‌دهد. طول این جاده ۱۸۰ کیلومتر که ۸۰ کیلومتر از این جاده یک‌طرفه و مابقی آن دوطرفه می‌باشد و به خاطر وجود پالایشگاه گاز خانگیران سرخس از اهمیت بسیار برخوردار می‌باشد. این مسیر دارای دو خط عبور می‌باشد.



شکل ۴-۱ پلان مسیر مشهد – سرخس

مسیر مشهد – چناران

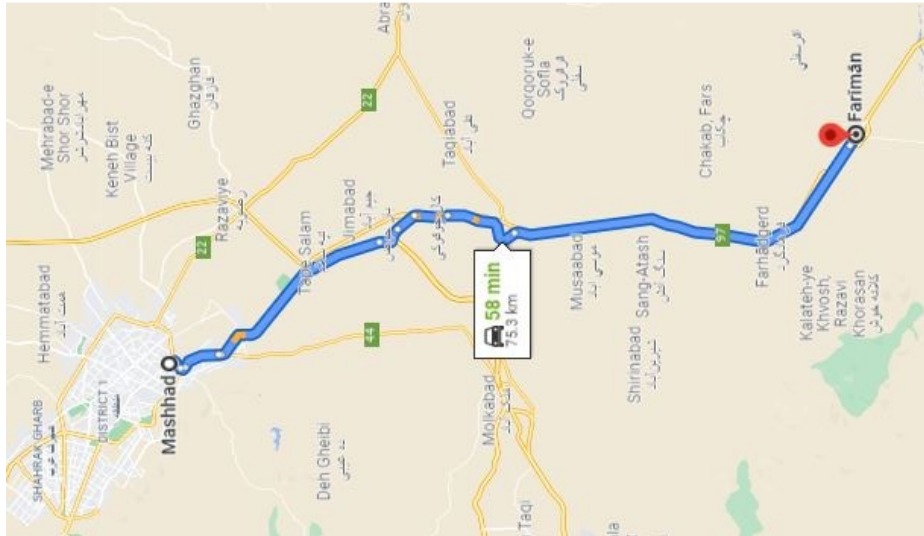
این جاده جزء پرترددترین جاده‌های منتهی به مشهد می‌باشد که ارتباط شبکه‌ای بین مسیر شمال کشور، کمربندی مشهد، جاده مشهد-ملک‌آباد و آزادراه حرم تا حرم را ایجاد می‌کند. طول این بزرگراه در حدود ۴۵ کیلومتر و اغلب مسیر دارای دو خط عبور می‌باشد.



شکل ۴-۲ پلان مسیر مشهد – چناران

مسیر مشهد - فریمان

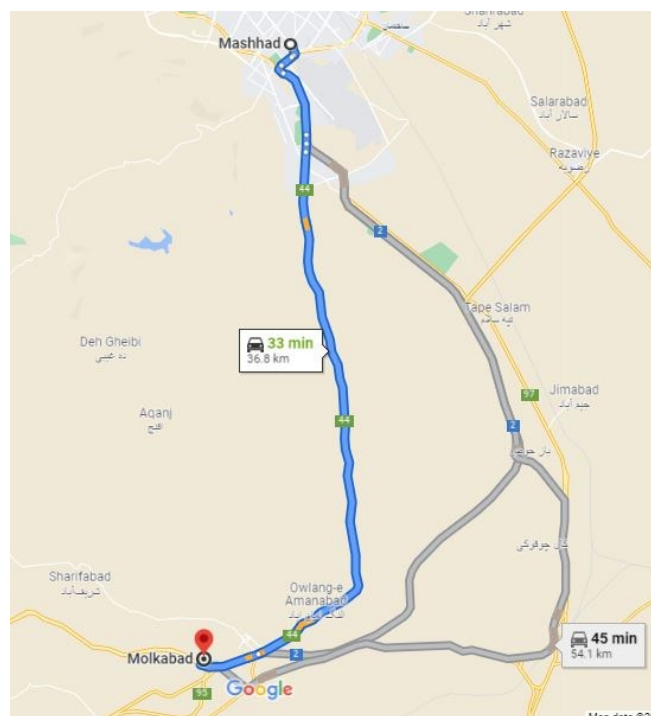
این جاده به طول ۶۵ کیلومتر است و این جاده مشهد را به کشور افغانستان ارتباط می دهد. و جزء جاده های ترانزیتی استان خراسان رضوی محسوب می شود. و اغلب مسیر دارای دو عبور می باشد.



شکل ۳-۴ پلان مسیر مشهد - فریمان

مسیر مشهد - ملک آباد

این مسیر به طول ۳۵ کیلومتر مشهد را به نیشابور در مسیر تهران و همچنین تربت حیدریه در مسیر زاهدان ارتباط می دهد. آزادراه راه مشهد - ملک آباد دارای دو باند مجزا و سه خط عبوری می باشد.



شکل ۴-۴ پلان مسیر مشهد - ملک آباد

مسیر مشهد – سدکارده

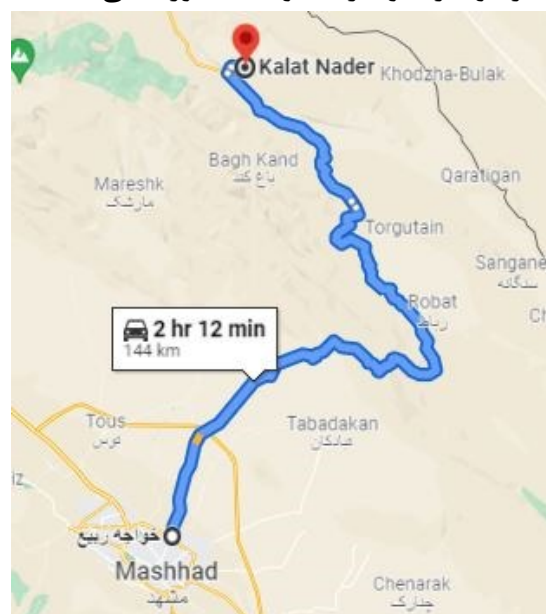
این جاده جزء جاده‌های توریستی-تفریحی محسوب می‌شود و از حجم تردد نسبتاً کمی برخوردار است. این مسیر یک‌طرفه و دارای دو خط عبور می‌باشد.



شکل ۴-۵ پلان مسیر مشهد – سدکارده

مسیر مشهد – کلات

این مسیر نیز مانند مسیر مشهد-سدکارده توریستی-تفریحی محسوب می‌شود و از حجم تردد کمی برخوردار است. این مسیر دو طرفه و دارای دو خط عبوری می‌باشد.



شکل ۴-۶ پلان مسیر مشهد – کلات

۴ - ۳ گردآوری داده ها

در پژوهش حاضر، با استفاده از سازمان راهداری و حمل و نقل جاده‌ای کشور داده‌های روزانه برای مسیرهای مورد نظر استخراج شده‌است. با توجه به داده‌های استخراج شده اولین روز برداشت داده مسیرها با هم متفاوت است (دلیل این امر می تواند زمان اجرا روش برخط در هر محور باشد). از این رو جدول ۱-۴ ساختار کلی مسیرهای موجود معرفی و باز زمانی داده‌های برداشت شده را نمایش می دهد. از آنجایی که داده‌های سامانه ترددشمار به دلایل مختلف در برخی از ایام سال برداشت نشده‌اند در پژوهش حاضر برای تخمین حجم تردد این ایام از روش میانگین‌گیری از روزهای سال قبل و بعد استفاده می‌گردد.

جدول ۱-۴ معرفی و بازه زمانی برداشت داده‌های مسیرهای استان خراسان رضوی

نام مسیر	نوع راه	تعداد خطوط عبور	آغاز برداشت	اتمام برداشت
مشهد - سرخس	بزرگراه	۴	۱۳۸۹/۱۲/۰۱	۱۳۹۸/۱۲/۲۹
مشهد - چناران	بزرگراه	۴-۶	۱۳۸۹/۰۱/۰۱	
مشهد - فریمان	بزرگراه	۴	۱۳۸۹/۱۲/۰۱	
مشهد - ملک آباد	آزادراه	۶	۱۳۸۹/۰۱/۰۱	
مشهد - سد کارده	راه اصلی	۲	۱۳۹۰/۰۶/۰۱	
مشهد - کلات	بزرگراه	۴	۱۳۹۱/۰۱/۰۱	

جدول ۲-۴ اطلاعات کلی از داده‌های مورد ارزیابی و پیش‌بینی

مسیر	روزهای برداشت	Max	زمان max	Min	زمان min
مشهد - سرخس	۳۳۱۶	۹۸۹۶	۱۳۹۷/۰۵/۱۹	۱۸۲۲	۱۳۹۱/۰۹/۰۴
سرخس - مشهد		۹۸۰۸	۱۳۹۲/۰۷/۳۰	۲۲۳	۱۳۹۱/۱۰/۰۷
مشهد - چناران	۳۶۵۲	۶۵۴۷۰	۱۳۹۷/۰۶/۱۹	۴۵۴۴	۱۳۹۰/۰۱/۰۱
چناران - مشهد		۶۳۴۶۰	۱۳۹۸/۰۳/۱۵	۲۰۶۹	۱۳۹۲/۱۰/۱۲
مشهد - فریمان	۳۳۱۶	۱۵۰۴۴	۱۳۹۱/۰۳/۲۲	۶۸۸	۱۳۹۱/۱۱/۰۶
فریمان - مشهد		۱۴۸۱۱	۱۳۹۸/۰۳/۱۷	۵۷	۱۳۹۰/۰۶/۰۷
مشهد - ملک آباد	۳۶۵۲	۵۶۶۵۹	۱۳۹۸/۰۸/۰۴	۲۲۳۳	۱۳۹۲/۱۲/۱۳
ملک آباد - مشهد		۶۶۳۰۲	۱۳۹۷/۰۸/۱۵	۹۶۵	۱۳۹۰/۱۱/۰۹
مشهد - سدکارده	۳۱۳۲	۱۱۳۳۹	۱۳۹۴/۰۱/۱۳	۵۶	۱۳۹۰/۱۰/۰۴
سدکارده - مشهد		۱۰۸۴۶	۱۳۹۲/۰۱/۱۳	۲۶	۱۳۹۴/۱۱/۰۴
مشهد - کلات	۲۹۵۴	۲۰۷۰۵	۱۳۹۴/۰۱/۱۳	۷۳	۱۳۹۴/۰۲/۱۸
کلات - مشهد		۱۵۰۴۶	۱۳۹۴/۰۱/۱۳	۶۵۳	۱۳۹۳/۰۱/۰۶

در جدول ۴-۲ برای ۱۲ مسیر موجود اطلاعات کلی از داده‌های این مسیرها از قبیل تعداد کل روزهای برداشت داده، ماکزیمم و مینیمم حجم تردد جاده‌ای در یک روز مشخص شده‌است.

۴ - ۴ مدل‌سازی و پیش‌بینی به روش ARIMA

در اولین قدم داده‌های حجم تردد روزانه را در یک ستون از نرم افزار SPSS بارگذاری کرده و تعیین کرده که اولین روز مربوط به چه روزی (شنبه، یکشنبه، دوشنبه، سه‌شنبه، چهارشنبه، پنج‌شنبه و یا جمعه) می‌باشد و هفته اول برداشت به عنوان اولین هفته برای نرم افزار تعریف می‌شود.

در قدم بعدی به رسم نمودار داده‌ها برای تشخیص نوع مدل پرداخته می‌شود. در قدم بعدی با استفاده از نمودار توابع ACF و PACF به بررسی ایستایی در میانگین پرداخته می‌شود. اگر این نمودارها پایداری در میانگین را بیان نکنند با تفاضل گیری درجه اول ($d=1$) و بررسی نمودارهای یک بار تفاضلی شده توابع ACF و PACF به بررسی دوباره ایستایی در میانگین پرداخته می‌شود (لازم به ذکر است در تمامی مسیرهای موجود حداکثر تا یکبار تفاضلی، ایستایی در میانگین محقق شده‌است). مرحله تشخیص و برازش مدل مولفه‌های p ، d و q برای شاخص غیرفصلی و P ، D و Q برای شاخص فصلی تشخیص داده می‌شوند و به برآورد پارامترهای مجهول پرداخته می‌شود.

بررسی مناسب بودن مدل با سه روش انجام می‌شود:

۱. **بررسی فرض نرمال بودن باقیمانده‌ها:** در صورتی که مدل انتخابی درست تشخیص داده شود باید خطاها دارای توزیع نرمال، با میانگین صفر و واریانس یک باشند. در صورتی که توزیع خطاها نرمال باشد در نمودار احتمال نرمال باید نقاط در امتداد یک خط مستقیم قرار گیرند.

۲. **بررسی فرض استقلال باقیمانده‌ها:** برای بررسی استقلال یا تصادفی بودن باقیمانده‌ها می‌توان از نمودارهای ACF و PACF باقیمانده‌ها استفاده کرد. چنانچه نمودارها روند خاصی را نشان ندهند و از حدود مجاز تجاوز نکنند می‌توان استقلال باقیمانده‌ها را پذیرفت.

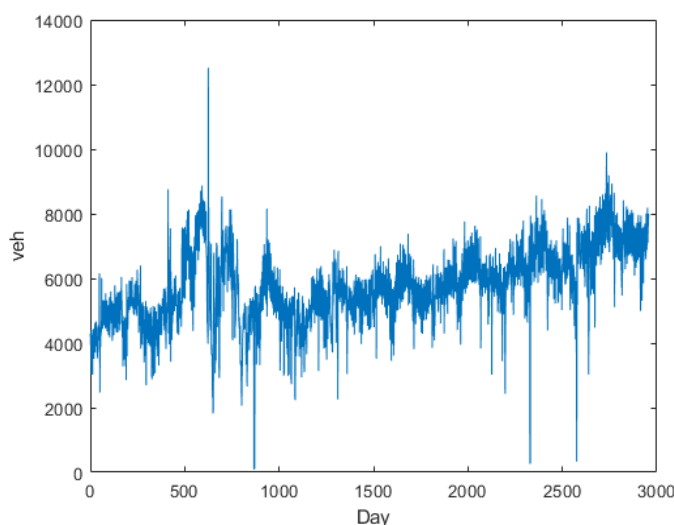
۳. **بررسی فرض ثابت بودن واریانس باقیمانده‌ها:** نمودار باقیمانده‌ها را در برابر زمان رسم می‌شود. در صورت ثابت بودن واریانس باقیمانده‌ها این نمودار باید فاقد ساختار باشد.

باتوجه به اشکال مورد نظر در تمامی مسیرهای موجود، فرض‌های بیان شده تایید می‌شوند و می‌توان به پیش‌بینی نهایی پرداخت و داده‌ها را برای یکسال آینده (سال ۱۳۹۸) به صورت روزانه پیش‌بینی نمود.

شکل‌های مربوط به نمودارهای ACF و PACF مربوط به داده‌های اصلی و یکبار تفاضلی شده و همچنین اشکال مربوط به بخش بررسی مناسب بودن مدل در پیوست مطالب ارائه شده‌اند.

مسیر مشهد - سرخس

شکل ۴-۷ حجم تردد روزانه برای ۳۳۱۶ روز در محور مشهد - سرخس مشخص شده است.



شکل ۴-۷ حجم تردد روزانه مسیر مشهد - سرخس

در شکل ۴-۷ محور عمودی بیانگر تعداد خودرو عبوری در محور مورد نظر و محور افقی تعداد روزهای برداشت داده‌های حجم تردد می‌باشد. برای داده‌های حجم تردد روزانه جاده‌ای در محور مشهد - سرخس مدل به صورت $ARIMA(2,1,1)(1,0,1)$ به دست آمده است (جدول ۴-۳).

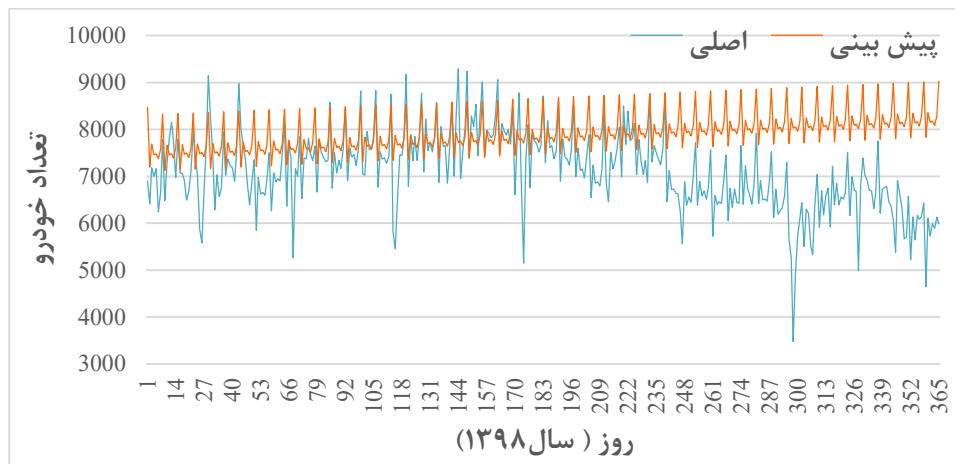
جدول ۴-۳ پارامترهای مدل برای محور مشهد - سرخس با حذف ضریب ثابت

پارامترها	تاخیر	مولفه‌ها	تخمین	SE	t	Sig.
AR	Lag ۱	p=۱	.۳۵۴	.۰۲۳	۱۵,۱۲۴	.۰۰۰
	Lag ۲	p=۲	.۰۷۲	.۰۲۱	۳,۳۵۵	.۰۰۱
Difference		d=۱	۱			
MA	Lag ۱	q=۱	.۸۸۳	.۰۱۴	۶۳,۵۰۵	.۰۰۰
AR, Seasonal	Lag ۱	P=۱	۱,۰۰۰	.۰۰۰	۵۰۴۲,۳۹	.۰۰۰
					۱	
MA, Seasonal	Lag ۱	Q=۱	.۹۸۶	.۰۰۵	۲۱۸,۶۳۸	.۰۰۰

در جدول ۴-۳ SE بیانگر انحراف از معیار، t بیانگر آماره t و sig. بیانگر p-value می‌باشند. با توجه به خروجی جدول، چون مقدار معناداری یعنی Sig. کوچکتر از سطح آزمون، یعنی ۰/۰۵ شده است، فرضیه صفر را رد و فرض صفر بودن مقدار ضریب ثابت مدل تایید می‌شود.

شکل ۴-۸ نمودار تغییرات حجم تردد روزانه نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که از شکل مشخص می‌باشد داده‌های اصلی در فصل بهار و تابستان تقریباً از یک روند صعودی و خطی پیروی می‌کنند که روش ARIMA این روند را تشخیص داده است و پیش‌بینی

قابل قبولی را ارائه می‌دهد اما برای داده‌های دیگر که غیرخطی هستند پیش‌بینی قابل قبولی را ارائه نمی‌دهد.



شکل ۴-۸ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - سرخس

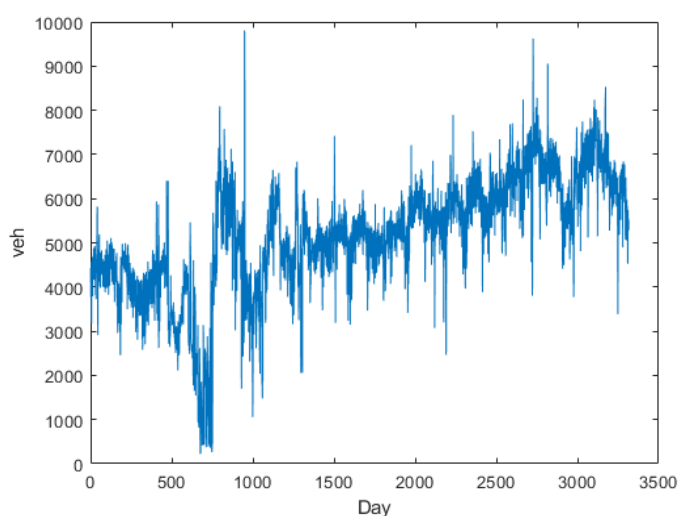
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۴ آمده است.

جدول ۴-۴ مقادیر خطا در محور مشهد - سرخس

شاخص	خطا
MSE	۰,۳۸۴
RMSE	۰,۶۱۹
MAPE%	۱۴,۳

مسیر سرخس - مشهد

شکل ۴-۹ حجم تردد روزانه برای ۳۳۱۶ روز در محور سرخس - مشهد مشخص شده است.



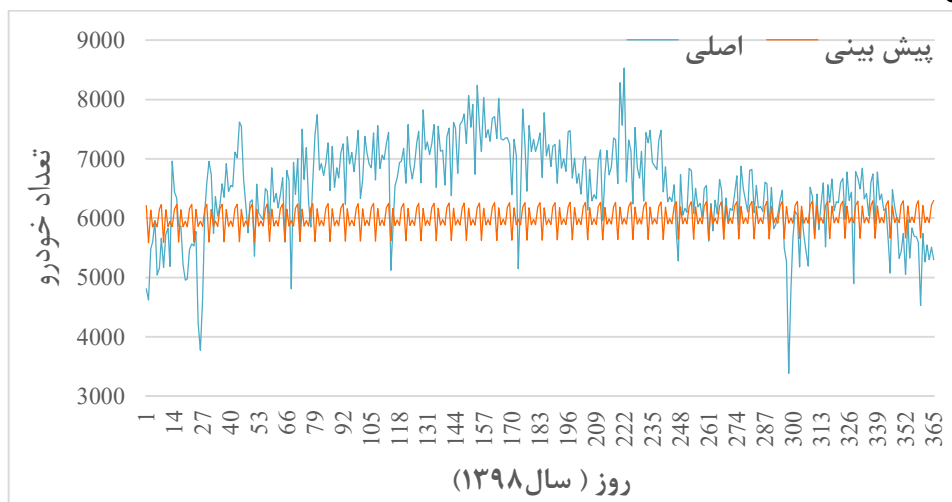
شکل ۴-۹ حجم تردد روزانه مسیر سرخس - مشهد

برای داده‌های حجم تردد روزانه جاده‌ای در محور سرخس - مشهد مدل به صورت $ARIMA(2,1,1)(1,0,1)$ به دست آمده است (جدول ۴-۵).

جدول ۴-۵ پارامترهای مدل برای محور سرخس - مشهد با حذف ضریب ثابت

پارامترها	تاخیر	مولفه‌ها	تخمین	SE	t	Sig.
AR	Lag ۱	p=۱	۰,۳۱۴	۰,۰۲۴	۱۳,۱۱۷	۰,۰۰۰
	Lag ۲	p=۲	۰,۰۴۳	۰,۰۲۲	۱,۹۹۲	۰,۰۴۶
Difference		d=۱	۱			
MA	Lag ۱	q=۱	۰,۸۶۵	۰,۰۱۵	۵۷,۹۸۵	۰,۰۰۰
AR, Seasonal	Lag ۱	P=۱	۱,۰۰۰	۰,۰۰۰	۳۹۵۵,۲۷	۰,۰۰۰
					۸	
MA, Seasonal	Lag ۱	Q=۱	۰,۹۹۱	۰,۰۰۵	۲۱۶,۲۰۸	۰,۰۰۰

در جدول ۴-۵ با توجه به خروجی جدول، چون مقدار معناداری یعنی Sig. کوچکتر از سطح آزمون، یعنی ۰/۰۵ شده است، فرضیه صفر را رد و فرض صفر بودن مقدار ضریب ثابت مدل تایید می‌شود. شکل ۴-۱۰ نمودار تغییرات حجم تردد روزانه نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد.



شکل ۴-۱۰ نمودار تغییرات حجم تردد نسبت به زمان در محور سرخس - مشهد

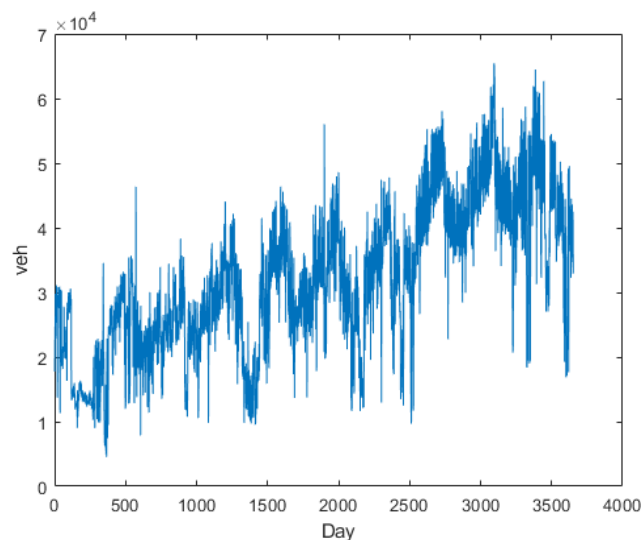
همانطور که از شکل ۴-۱۰ مشخص می‌باشد داده‌های اصلی در فصل بهار و زمستان تقریباً از یک روند صعودی و خطی نسبت به داده‌های آموزش دارند پیروی می‌کنند که روش $ARIMA$ این روند را تشخیص داده است و پیش‌بینی قابل قبولی را ارائه می‌دهد اما برای داده‌های دیگر که حجم تردد نسبت به سال‌های گذشته رشد محسوسی داشته است پیش‌بینی قابل قبولی را ارائه نمی‌دهد. مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۶ آمده است.

جدول ۴-۶ مقادیر خطا در محور سرخس - مشهد

خطا	شاخص
MSE	۰,۳۲۳
RMSE	۰,۵۶۹
MAPE%	۱۱,۳

مسیر مشهد - چناران

شکل ۴-۱۱ حجم تردد روزانه برای ۳۶۵۲ روز در محور مشهد - چناران مشخص شده است.



شکل ۴-۱۱ حجم تردد روزانه مسیر مشهد - چناران

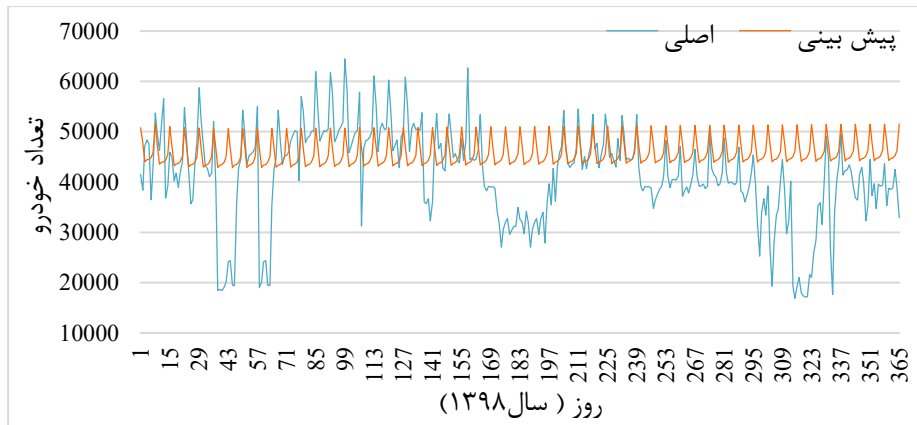
برای داده‌های حجم تردد روزانه جاده‌ای در مسیر مشهد - چناران مدل به صورت $ARIMA(2,1,2)(1,0,1)$ به دست آمده است (جدول ۴-۷).

جدول ۴-۷ پارامترهای مدل برای مسیر مشهد - چناران با حذف ضریب ثابت

پارامترها	تاخیر	مولفه‌ها	تخمین	SE	t	Sig.
AR	Lag ۱	p=۱	۱,۱۹۹	۰,۰۷۱	۱۶,۸۵۵	۰,۰۰۰
	Lag ۲	p=۲	-۰,۲۵۵	۰,۰۴۹	-۵,۲۴۳	۰,۰۰۰
Difference		d=۱	۱			
MA	Lag ۱	q=۱	۱,۶۲۸	۰,۰۶۶	۲۴,۷۰۳	۰,۰۰۰
	Lag ۲	q=۲	-۰,۶۳۸	۰,۰۶۰	-۱۰,۶۱۴	۰,۰۰۰
AR, Seasonal	Lag ۱	P=۱	۰,۹۹۹	۰,۰۰۱	۸۱۵,۵۳۸	۰,۰۰۰
MA, Seasonal	Lag ۱	Q=۱	۰,۹۷۰	۰,۰۰۶	۱۶۴,۳۱۳	۰,۰۰۰

در جدول ۴-۶ با توجه به خروجی جدول، چون مقدار معناداری یعنی Sig. کوچکتر از سطح آزمون، یعنی ۰/۰۵ شده است، فرضیه صفر را رد و فرض صفر بودن مقدار ضریب ثابت مدل تایید می‌شود.

شکل ۴-۱۲ نمودار تغییرات حجم تردد روزانه نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که از شکل مشخص می‌باشد داده‌های اصلی از یک روند خاصی پیروی نمی‌کنند و خطی نیستند. پس مدل ARIMA نمی‌تواند پیش‌بینی قابل قبولی را برای این سری از داده‌ها ارائه دهد.



شکل ۴-۱۲ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - چناران

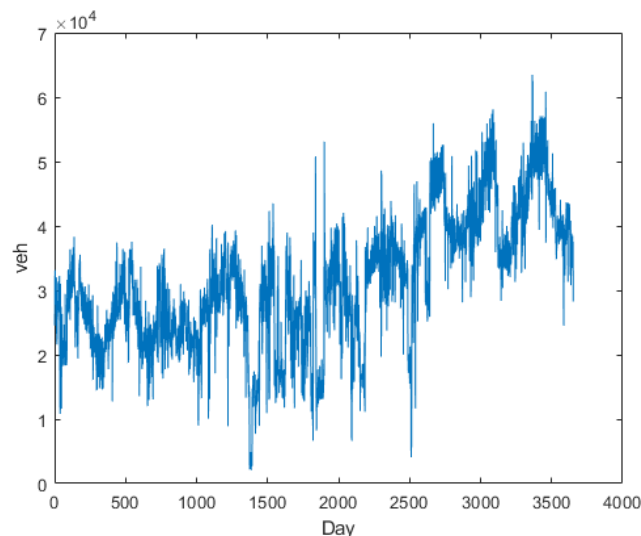
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۸ آمده است.

جدول ۴-۸ مقادیر خطا در محور مشهد - چناران

شاخص	خطا
MSE	۰,۵۴۴
RMSE	۰,۷۳۸
MAPE%	۲۴,۴

مسیر چناران - مشهد

شکل ۴-۱۳ حجم تردد روزانه برای ۳۶۵۲ روز در محور چناران - مشهد مشخص شده است.



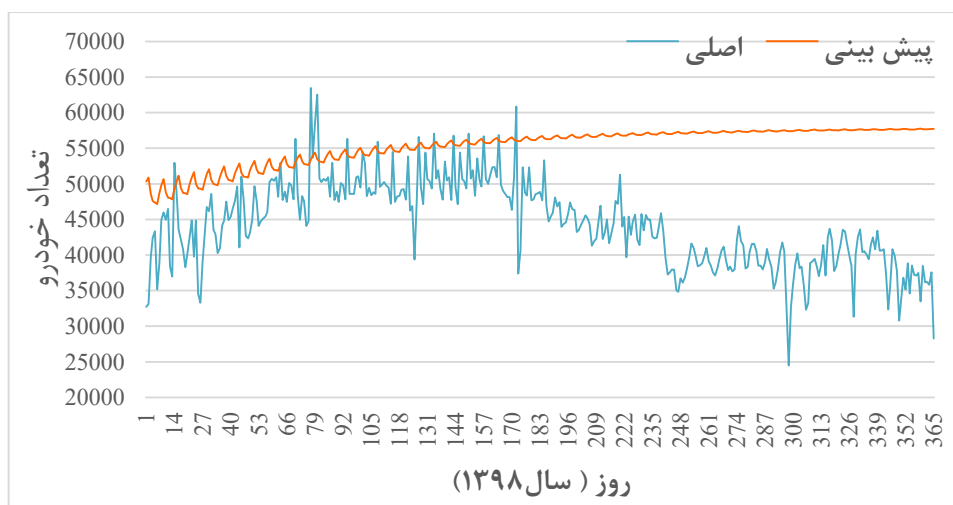
شکل ۴-۱۳ حجم تردد روزانه مسیر چناران - مشهد

برای داده‌های حجم تردد روزانه جاده‌ای در مسیر چناران - مشهد مدل به صورت $ARIMA(2,1,1)(1,0,1)$ به دست آمده است (جدول ۹-۴).

جدول ۹-۴ پارامترهای مدل برای محور چناران - مشهد با حذف ضریب ثابت

پارامترها	تاخیر	مولفه‌ها	تخمین	SE	t	Sig.
AR	Lag ۱	p=۱	۰,۴۰۳	۰,۰۲۷	۱۴,۶۸۶	۰,۰۰۰
	Lag ۲	p=۲	۰,۱۰۸	۰,۰۲۲	۴,۸۴۸	۰,۰۰۰
Difference		d=۱	۱			
MA	Lag ۱	q=۱	۰,۸۵۱	۰,۰۲۰	۴۲,۰۴۸	۰,۰۰۰
AR, Seasonal	Lag ۱	P=۱	۰,۹۳۱	۰,۰۱۴	۶۴,۲۷۵	۰,۰۰۰
MA, Seasonal	Lag ۱	Q=۱	۰,۸۲۶	۰,۰۲۳	۳۶,۳۷۴	۰,۰۰۰

در جدول ۹-۴ با توجه به خروجی جدول، چون مقدار معناداری یعنی Sig. کوچکتر از سطح آزمون، یعنی ۰/۰۵ شده است، فرضیه صفر را رد و فرض صفر بودن مقدار ضریب ثابت مدل تایید می‌شود. شکل ۱۴-۴ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد.



شکل ۱۴-۴ نمودار تغییرات حجم تردد نسبت به زمان در محور چناران - مشهد

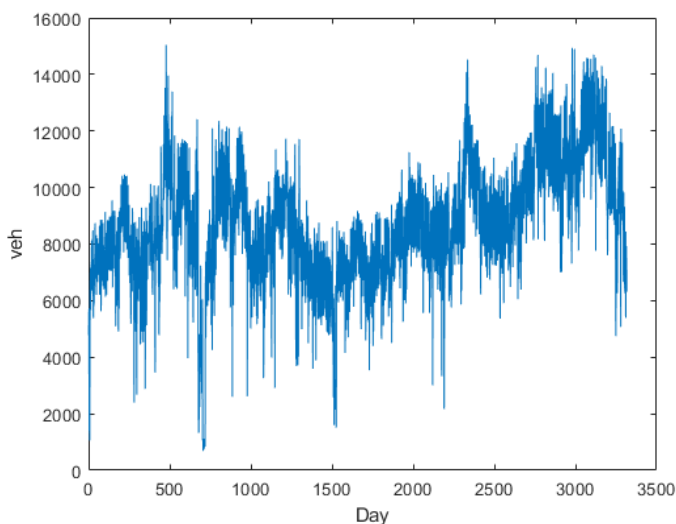
همانطور که از شکل ۱۴-۴ مشخص می‌باشد داده‌های اصلی در فصل بهار و تابستان تقریباً از یک روند صعودی و خطی پیروی می‌کنند که روش $ARIMA$ این روند را تشخیص داده است و پیش‌بینی قابل قبولی را ارائه می‌دهد اما برای داده‌های دیگر که غیرخطی هستند پیش‌بینی قابل قبولی را ارائه نمی‌دهد. مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۱۰-۴ آمده است.

جدول ۴-۱۰ مقادیر خطا در محور چناران - مشهد

خطا	شاخص
۰,۶۹	MSE
۰,۸۳۱	RMSE
۲۸	MAPE%

مسیر مشهد - فریمان

شکل ۴-۱۵ حجم تردد روزانه برای ۳۳۱۶ روز در محور مشهد - فریمان مشخص شده است.



شکل ۴-۱۵ حجم تردد روزانه مسیر مشهد - فریمان

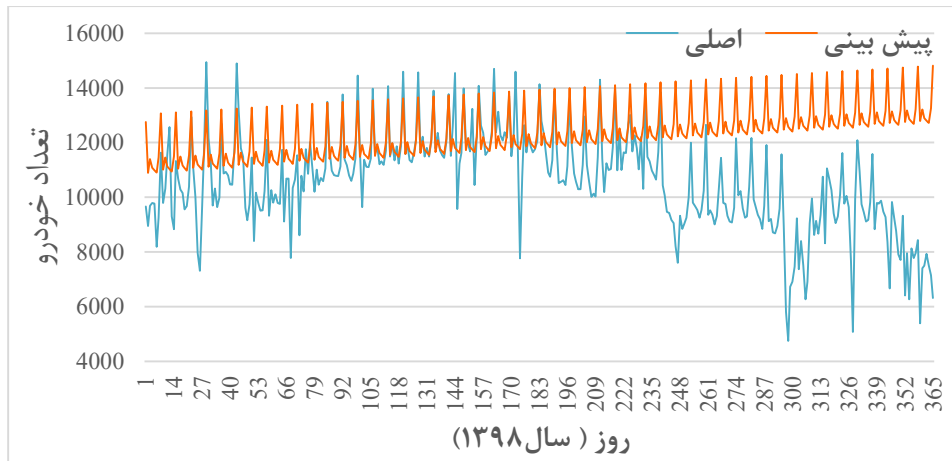
برای داده‌های حجم تردد روزانه جاده‌ای در مسیر مشهد - فریمان مدل به صورت $ARIMA(1,1,1)(10,1)$ به دست آمده است (جدول ۴-۱۱).

جدول ۴-۱۱ پارامترهای مدل برای مسیر مشهد - فریمان با حذف ضریب ثابت

پارامترها	تاخیر	مولفه‌ها	تخمین	SE	t	Sig.
AR	Lag ۱	p=۱	۰,۲۷۶	۰,۰۲۵	۱۰,۹۲۹	۰,۰۰۰
Difference		d=۱	۱			
MA	Lag ۱	q=۱	۰,۸۲۰	۰,۰۱۵	۵۴,۵۴۴	۰,۰۰۰
AR, Seasonal	Lag ۱	P=۱	۰,۹۹۹	۰,۰۰۰	۲۱۱۸,۸۷ ۶	۰,۰۰۰
MA, Seasonal	Lag ۱	Q=۱	۰,۹۷۳	۰,۰۰۶	۱۷۰,۰۲۲	۰,۰۰۰

در جدول ۴-۱۱ با توجه به خروجی جدول، چون مقدار معناداری یعنی Sig. کوچکتر از سطح آزمون، یعنی ۰/۰۵ شده است، فرضیه صفر را رد و فرض صفر بودن مقدار ضریب ثابت مدل تایید می‌شود.

شکل ۴-۱۶ نمودار تغییرات حجم تردد روزانه نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که از شکل مشخص می‌باشد داده‌های اصلی از یک روند خاصی پیروی نمی‌کنند و خطی نیستند. پس مدل ARIMA نمی‌تواند پیش‌بینی قابل قبولی را برای این سری از داده‌ها ارائه دهد.



شکل ۴-۱۶ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - فریمان

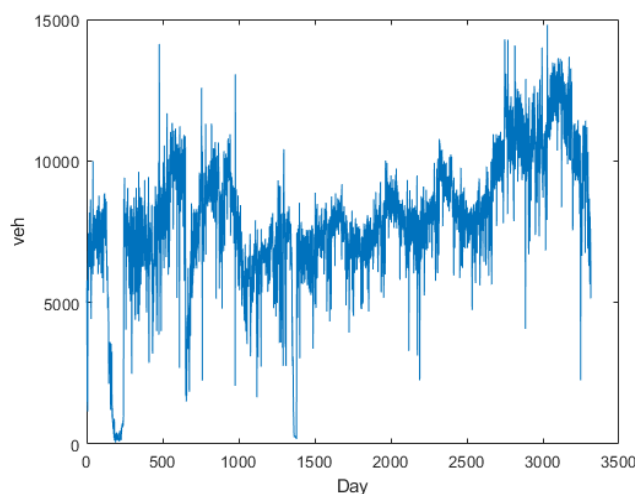
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۱۲ آمده است.

جدول ۴-۱۲ مقادیر خطا در محور مشهد - فریمان

شاخص	خطا
MSE	۰,۴۶۴
RMSE	۰,۶۸۱
MAPE%	۲۱,۵

مسیر فریمان - مشهد

شکل ۴-۱۷ حجم تردد روزانه برای ۳۳۱۶ روز در محور فریمان - مشهد مشخص شده است.



شکل ۴-۱۷ حجم تردد روزانه مسیر فریمان - مشهد

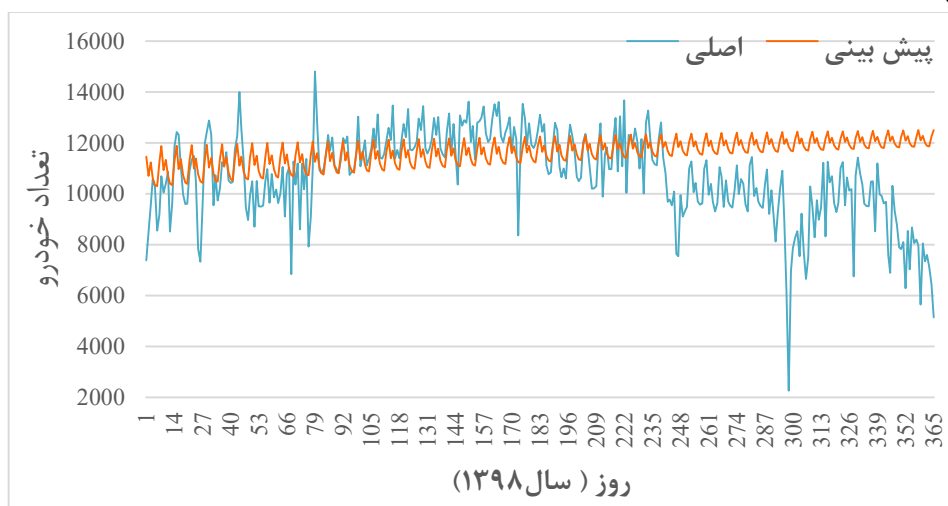
برای داده‌های حجم تردد روزانه جاده‌ای در محور فریمان - مشهد مدل به صورت $ARIMA(3,1,2)(1,0,1)$ به دست آمده است (جدول ۴-۱۳).

جدول ۴-۱۳ پارامترهای مدل برای محور فریمان - مشهد با حذف ضریب ثابت

پارامترها	تاخیر	مولفه‌ها	تخمین	SE	t	Sig.
AR	Lag ۱	p=۱	-.۰۷۵۱	۰.۰۲۹	-۲۶.۳۴۳	۰.۰۰۰
	Lag ۲	p=۲	۰.۳۰۵	۰.۰۴۳	۷.۱۵۱	۰.۰۰۰
	Lag ۳	p=۳	۰.۰۵۷	۰.۰۲۴	۲.۴۲۷	۰.۰۱۵
Difference		d=۱	۱			
MA	Lag ۱	q=۱	-.۰۱۸۷	۰.۰۲۱	-۸.۸۷۰	۰.۰۰۰
	Lag ۲	q=۲	۰.۸۰۶	۰.۰۲۱	۳۸.۳۶۳	۰.۰۰۰
AR, Seasonal	Lag ۱	P=۱	۰.۹۸۳	۰.۰۰۵	۱۸۲.۱۱۶	۰.۰۰۰
MA, Seasonal	Lag ۱	Q=۱	۰.۹۱۴	۰.۰۱۲	۷۷.۱۰۷	۰.۰۰۰

در جدول ۴-۱۳ با توجه به خروجی جدول، چون مقدار معناداری یعنی Sig. کوچکتر از سطح آزمون، یعنی ۰/۰۵ شده است، فرضیه صفر را رد و فرض صفر بودن مقدار ضریب ثابت مدل تایید می‌شود.

شکل ۴-۱۸ نمودار تغییرات حجم تردد روزانه نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد.



شکل ۴-۱۸ نمودار تغییرات حجم تردد نسبت به زمان در محور فریمان - مشهد

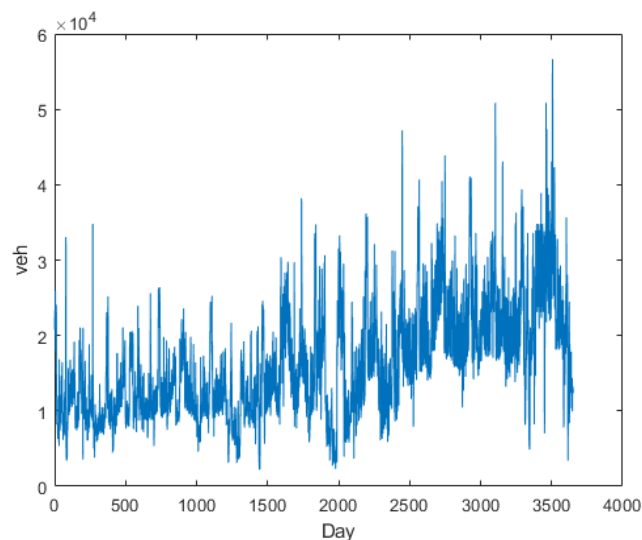
همانطور که از شکل ۴-۱۸ مشخص می‌باشد داده‌های اصلی در ۸ ماه اولیه سال تقریباً از یک روند صعودی و خطی پیروی می‌کنند که روش $ARIMA$ این روند را تشخیص داده است و پیش‌بینی قابل قبولی را ارائه می‌دهد اما برای داده‌های دیگر که غیرخطی هستند پیش‌بینی قابل قبولی را ارائه نمی‌دهد. مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۱۴ آمده است.

جدول ۴-۱۴ مقادیر خطا در محور فریمان - مشهد

شاخص	خطا
MSE	۰,۳۰۲
RMSE	۰,۵۴۹
MAPE%	۱۷

مسیر مشهد - ملک آباد

شکل ۴-۱۹ حجم تردد روزانه برای ۳۶۵۲ روز در محور مشهد - ملک آباد مشخص شده است.



شکل ۴-۱۹ حجم تردد روزانه مسیر مشهد - ملک آباد

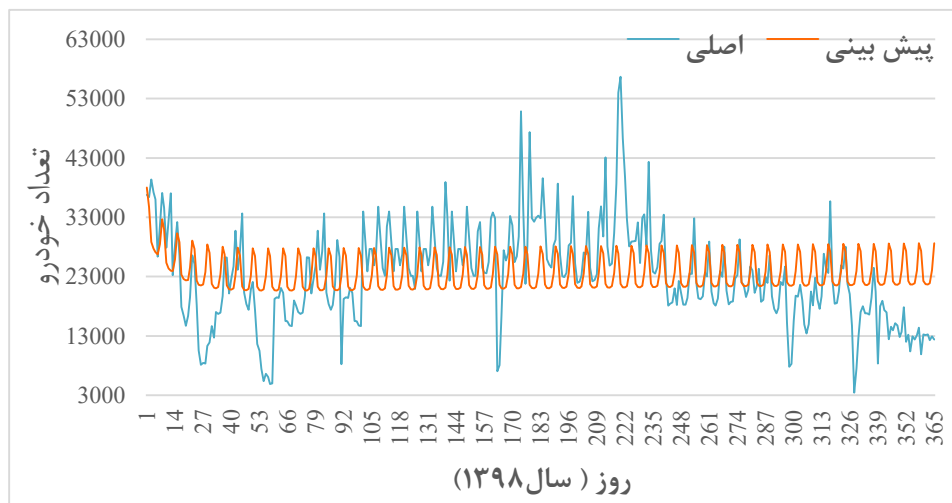
برای داده‌های حجم تردد روزانه جاده‌ای در مسیر مشهد - ملک آباد مدل به صورت $ARIMA(3,1,2)(1,0,1)$ به دست آمده است (جدول ۴-۱۵).

جدول ۴-۱۵ پارامترهای مدل برای مسیر مشهد - ملک آباد با حذف ضریب ثابت

پارامترها	تاخیر	مولفه‌ها	تخمین	SE	t	Sig.
AR	Lag ۱	p=۱	۱,۱۱۹	۰,۰۱۸	۶۳,۴۷۹	۰,۰۰۰
	Lag ۲	p=۲	-۰,۳۱۸	۰,۰۲۶	-۱۲,۳۹۰	۰,۰۰۰
	Lag ۳	p=۳	۰,۱۲۰	۰,۰۱۷	۶,۸۷۵	۰,۰۰۰
Difference		d=۱	۱			
MA	Lag ۱	q=۱	۱,۳۷۴	۰,۰۰۴	۳۸۲,۲۰۸	۰,۰۰۰
	Lag ۲	q=۲	-۰,۳۷۴	۰,۰۰۳	-۱۲۹,۷۷۱	۰,۰۰۰
AR, Seasonal	Lag ۱	P=۱	۰,۹۹۹	۰,۰۰۱	۱۸۳۸,۷۸	۰,۰۰۰
MA, Seasonal	Lag ۱	Q=۱	۰,۹۷۵	۰,۰۰۵	۱۹۰,۲۷۳	۰,۰۰۰

در جدول ۴-۱۵ با توجه به خروجی جدول، چون مقدار معناداری یعنی Sig. کوچکتر از سطح آزمون، یعنی ۰/۰۵ شده است، فرضیه صفر را رد و فرض صفر بودن مقدار ضریب ثابت مدل تایید می‌شود.

شکل ۴-۲۰ نمودار تغییرات حجم تردد روزانه نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که از شکل مشخص می‌باشد داده‌های اصلی از یک روند خاصی پیروی نمی‌کنند اما در فصل تابستان و ماه آذر و آبان به دلیل تغییرات نسبی کمتر نسبت به سال‌های قبل از پیش‌بینی قابل قبولی دارد



شکل ۴-۲۰ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - ملک آباد

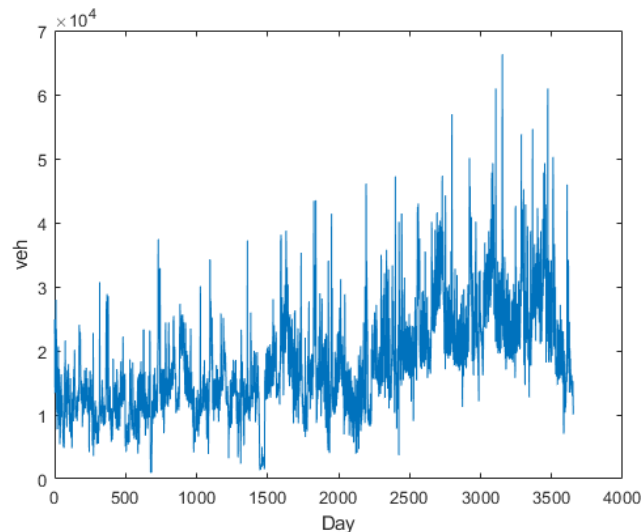
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۱۶ آمده است.

جدول ۴-۱۶ مقادیر خطا در محور مشهد - ملک آباد

شاخص	خطا
MSE	۰,۳۲۵
RMSE	۰,۵۷
MAPE%	۳۶,۶

مسیر ملک آباد - مشهد

شکل ۴-۲۱ حجم تردد روزانه برای ۳۶۵۲ روز در محور ملک آباد - مشهد مشخص شده است.



شکل ۴-۲۱ حجم تردد روزانه مسیر ملک آباد - مشهد

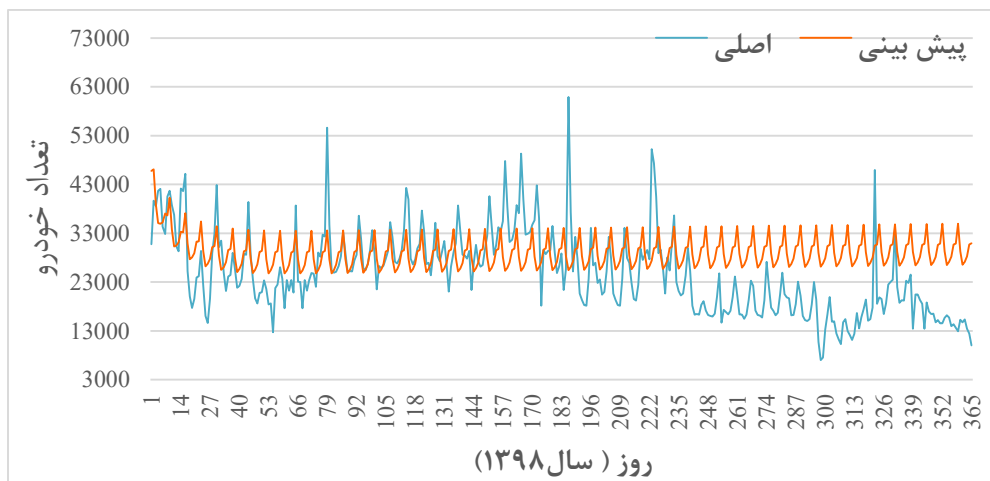
برای داده‌های حجم تردد روزانه جاده‌ای در محور ملک آباد - مشهد مدل به صورت $ARIMA(3,1,2)(1,0,1)$ به دست آمده است (جدول ۴-۱۷).

جدول ۴-۱۷ پارامترهای مدل برای محور ملک آباد - مشهد با حذف ضریب ثابت

پارامترها	تاخیر	مولفه‌ها	تخمین	SE	t	Sig.
AR	Lag ۱	p=۱	۱,۲۱۷	۰,۰۱۸	۶۷,۷۹۹	۰,۰۰۰
	Lag ۲	p=۲	-۰,۳۵۲	۰,۰۲۷	-۱۳,۰۶۴	۰,۰۰۰
	Lag ۳	p=۳	۰,۰۷۱	۰,۰۱۸	۴,۰۲۴	۰,۰۰۰
Difference		d=۱	۱			
MA	Lag ۱	q=۱	۱,۵۸۰	۰,۰۰۴	۳۷۴,۶۸۲	۰,۰۰۰
	Lag ۲	q=۲	-۰,۵۸۰	۰,۰۰۴	-۱۶۰,۷۲۸	۰,۰۰۰
AR, Seasonal	Lag ۱	P=۱	۰,۹۹۹	۰,۰۰۱	۱۱۰۸,۱۲۵	۰,۰۰۰
MA, Seasonal	Lag ۱	Q=۱	۰,۹۷۵	۰,۰۰۶	۱۶۸,۲۹۵	۰,۰۰۰

در جدول ۴-۱۷ با توجه به خروجی جدول، چون مقدار معناداری یعنی Sig. کوچکتر از سطح آزمون، یعنی ۰/۰۵ شده است، فرضیه صفر را رد و فرض صفر بودن مقدار ضریب ثابت مدل تایید می‌شود.

شکل ۴-۲۲ نمودار تغییرات حجم تردد روزانه نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد.



شکل ۴-۲۲ نمودار تغییرات حجم تردد نسبت به زمان در محور ملک آباد - مشهد

همانطور که از شکل ۴-۲۲ مشخص می‌باشد داده‌های اصلی در ۸ ماه اولیه سال تقریباً از یک روند صعودی و خطی پیروی می‌کنند که روش ARIMA این روند را تشخیص داده است و پیش‌بینی قابل قبولی را ارائه می‌دهد اما برای داده‌های دیگر که غیرخطی هستند پیش‌بینی قابل قبولی را ارائه نمی‌دهد.

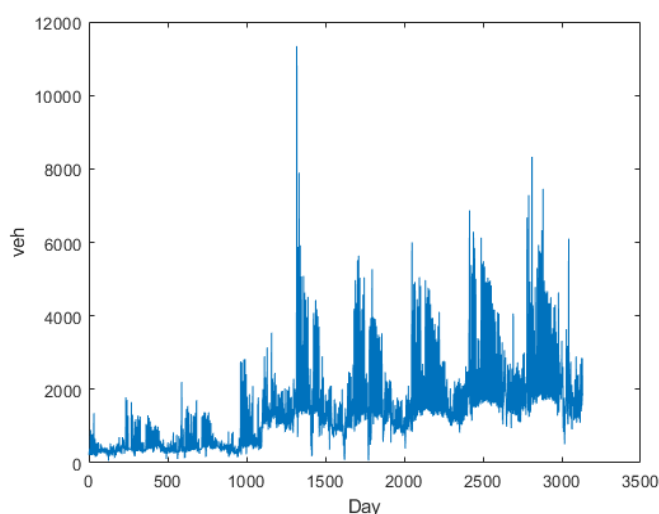
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۱۸ آمده است.

جدول ۴-۱۸ مقادیر خطا در محور ملک آباد - مشهد

خطا	شاخص
MSE	۰,۲۰۶
RMSE	۰,۴۵۴
MAPE%	۳۸,۸

مسیر مشهد - سدکارده

شکل ۴-۲۳ حجم تردد روزانه برای ۳۱۳۲ روز در محور مشهد - سدکارده مشخص شده است.



شکل ۴-۲۳ حجم تردد روزانه مسیر مشهد - سدکارده

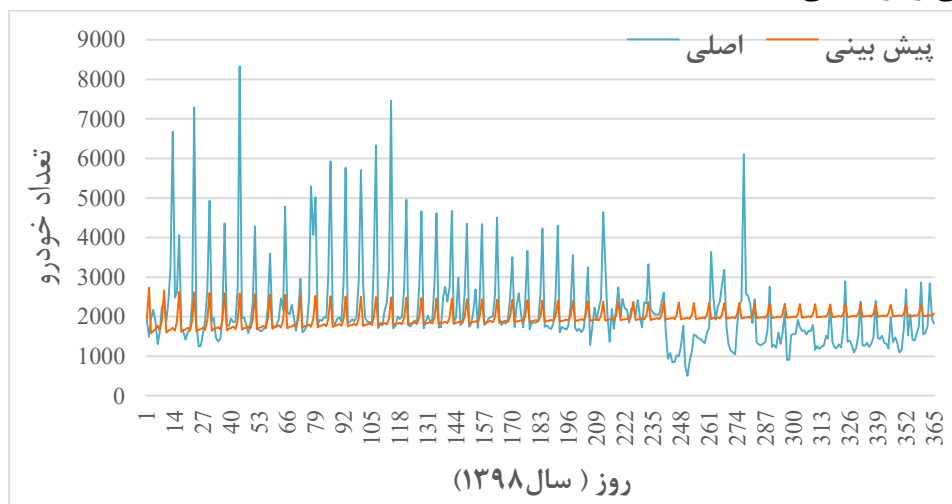
برای داده‌های حجم تردد روزانه جاده‌ای در مسیر مشهد - سدکارده مدل به صورت $ARIMA(1,1,1)(2,0,1)$ به دست آمده است (جدول ۴-۱۹).

جدول ۴-۱۹ پارامترهای مدل برای مسیر مشهد - سدکارده با حذف ضریب ثابت

پارامترها	تاخیر	مولفه‌ها	تخمین	SE	t	Sig.
AR	Lag ۱	p=۱	۰,۲۴۲	۰,۰۲۰	۱۲,۳۴۷	۰,۰۰۰
Difference		d=۱	۱			
MA	Lag ۱	q=۱	۰,۹۷۱	۰,۰۰۵	۱۸۶,۳۸۱	۰,۰۰۰
AR, Seasonal	Lag ۱	P=۱	۱,۱۷۰	۰,۰۲۹	۴۰,۵۲۹	۰,۰۰۰
	Lag ۲	P=۲	-۰,۱۹۱	۰,۰۲۶	-۷,۲۳۶	۰,۰۰۰
MA, Seasonal	Lag ۱	Q=۱	۰,۸۰۹	۰,۰۲۱	۳۹,۰۳۶	۰,۰۰۰

در جدول ۴-۱۹ با توجه به خروجی جدول، چون مقدار معناداری یعنی Sig. کوچکتر از سطح آزمون، یعنی ۰/۰۵ شده است، فرضیه صفر را رد و فرض صفر بودن مقدار ضریب ثابت مدل تایید می‌شود.

شکل ۴-۲۴ نمودار تغییرات حجم تردد روزانه نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که از شکل مشخص می‌باشد به دلیل کاربری خاص این مسیر (توریستی-تفریحی) داده‌های اصلی در روزهای غیر تعطیل از یک روند خاص و خطی پیروی می‌کنند که روش ARIMA این روند را تشخیص داده است و پیش‌بینی قابل قبولی را ارائه می‌دهد اما برای داده‌های آخر هفته و مناسب‌هایی که تعطیلی رسمی هستند داده‌ها تغییرات غیرخطی دارند و این مدل پیش‌بینی قابل قبولی را ارائه نمی‌دهد.



شکل ۴-۲۴ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - سدکارده

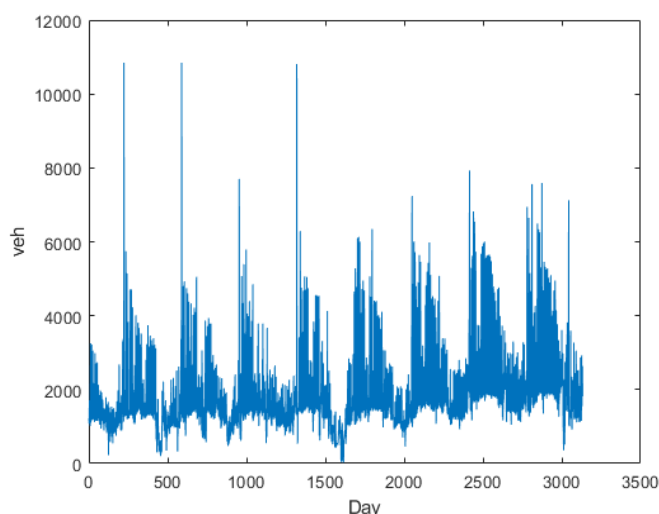
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۲۰ آمده است.

جدول ۴-۲۰ مقادیر خطا در محور مشهد - سدکارده

خطا	شاخص
۰,۱۸۲	MSE
۰,۴۲۶	RMSE
۰,۲۷۳	MAPE%

مسیر سدکارده - مشهد

شکل ۴-۲۵ حجم تردد روزانه برای ۳۱۳۲ روز در محور سدکارده - مشهد مشخص شده است.



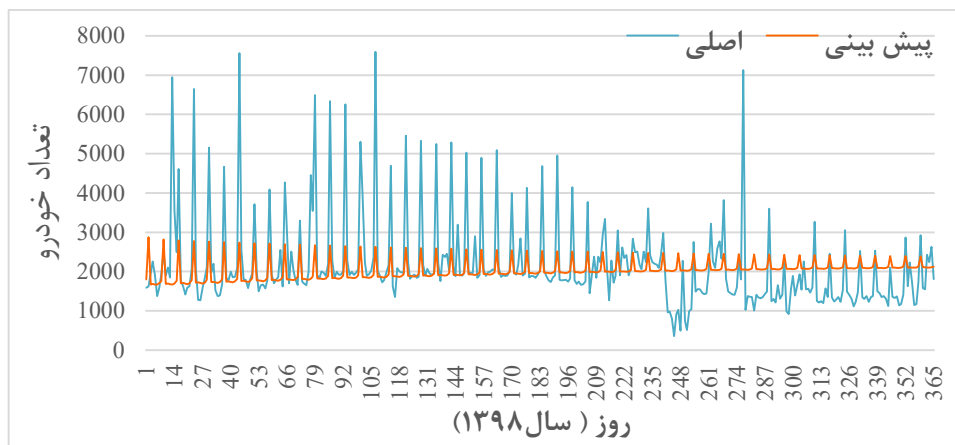
شکل ۴-۲۵ حجم تردد روزانه مسیر سدکارده - مشهد

برای داده‌های حجم تردد روزانه جاده‌ای در محور سدکارده - مشهد مدل به صورت $ARIMA(1,1,1)(2,0,1)$ به دست آمده است (جدول ۴-۲۱).

جدول ۴-۲۱ پارامترهای مدل برای محور سدکارده - مشهد با حذف ضریب ثابت

پارامترها	تاخیر	مولفه‌ها	تخمین	SE	t	Sig.
AR	Lag ۱	p=۱	۰,۱۵۳	۰,۰۲۰	۷,۶۶۹	۰,۰۰۰
Difference		d=۱	۱			
MA	Lag ۱	q=۱	۰,۹۶۵	۰,۰۰۶	۱۷۰,۶۰۸	۰,۰۰۰
AR, Seasonal	Lag ۱	P=۱	۱,۰۸۳	۰,۰۲۸	۳۸,۰۰۹	۰,۰۰۰
	Lag ۲	P=۲	-۰,۱۰۸	۰,۰۲۶	-۴,۱۵۴	۰,۰۰۰
MA, Seasonal	Lag ۱	Q=۱	۰,۷۸۸	۰,۰۲۱	۳۷,۶۸۸	۰,۰۰۰

در جدول ۴-۲۱ با توجه به خروجی جدول، چون مقدار معناداری یعنی Sig. کوچکتر از سطح آزمون، یعنی ۰/۰۵ شده است، فرضیه صفر را رد و فرض صفر بودن مقدار ضریب ثابت مدل تایید می‌شود. شکل ۴-۲۶ نمودار تغییرات حجم تردد روزانه نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد.



شکل ۴-۲۶ نمودار تغییرات حجم تردد نسبت به زمان در محور سدکارده - مشهد

همانطور که از شکل ۴-۲۶ مشخص می‌باشد به دلیل کاربری خاص این مسیر (توریستی - تفریحی) داده‌های اصلی در روزهای غیر تعطیل از یک روند خاص و خطی پیروی می‌کنند که روش ARIMA این روند را تشخیص داده است و پیش‌بینی قابل قبولی را ارائه می‌دهد اما برای داده‌های آخر، داده‌ها تغییرات غیر خطی دارند و این مدل پیش‌بینی قابل قبولی را ارائه نمی‌دهد.

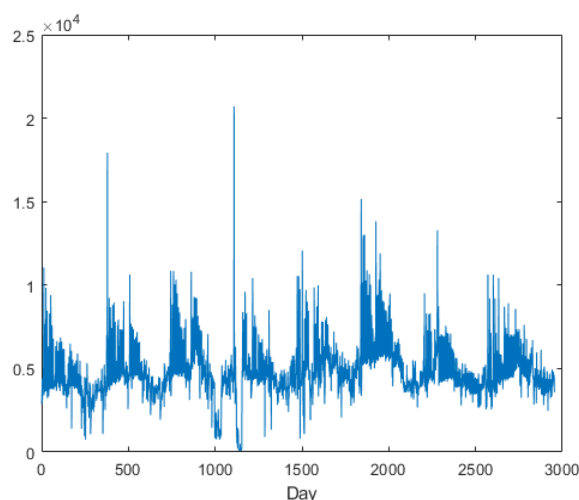
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۲۲ آمده است.

جدول ۴-۲۲ مقادیر خطا در محور سدکارده - مشهد

شاخص	خطا
MSE	۰,۱۳۵
RMSE	۰,۳۶۸
MAPE%	۲۹,۷

مسیر مشهد - کلات

شکل ۴-۲۷ حجم تردد روزانه برای ۲۹۵۴ روز در محور مشهد - کلات مشخص شده است.



شکل ۴-۲۷ حجم تردد روزانه مسیر مشهد - کلات

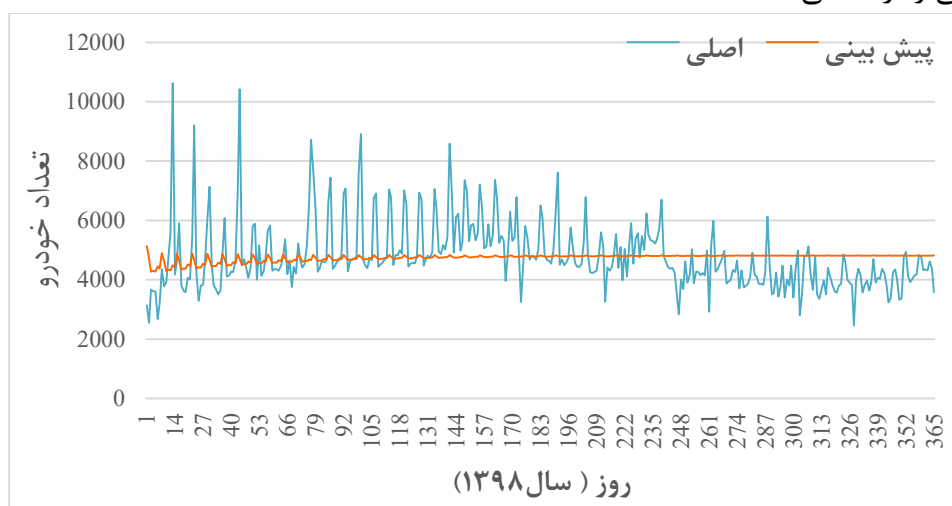
برای داده‌های حجم تردد روزانه جاده‌ای در مسیر مشهد - کلات مدل به صورت $ARIMA(1,1,1)(1,0,1)$ به دست آمده است (جدول ۴-۲۳).

جدول ۴-۲۳ پارامترهای مدل برای مسیر مشهد - کلات با حذف ضریب ثابت

پارامترها	تاخیر	مولفه‌ها	تخمین	SE	t	Sig.
AR	Lag ۱	p=۱	۰,۴۷۴	۰,۰۱۹	۲۴,۶۱۴	۰,۰۰۰
Difference		d=۱	۱			
MA	Lag ۱	q=۱	۰,۹۶۸	۰,۰۰۶	۱۶۰,۸۶۵	۰,۰۰۰
AR, Seasonal	Lag ۱	P=۱	۰,۹۰۰	۰,۰۱۴	۶۶,۴۲۷	۰,۰۰۰
MA, Seasonal	Lag ۱	Q=۱	۰,۶۲۰	۰,۰۲۶	۲۴,۲۰۵	۰,۰۰۰

در جدول ۴-۲۳ با توجه به خروجی جدول، چون مقدار معناداری یعنی Sig. کوچکتر از سطح آزمون، یعنی ۰/۰۵ شده است، فرضیه صفر را رد و فرض صفر بودن مقدار ضریب ثابت مدل تایید می‌شود.

شکل ۴-۲۸ نمودار تغییرات حجم تردد روزانه نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که از شکل مشخص می‌باشد به دلیل کاربری خاص این مسیر (توریستی-تفریحی) داده‌های اصلی در روزهای غیر تعطیل از یک روند خاص و خطی پیروی می‌کنند که روش ARIMA این روند را تشخیص داده است و پیش‌بینی قابل قبولی را ارائه می‌دهد اما برای داده‌های آخر هفته و مناسب‌هایی که تعطیلی رسمی هستند داده‌ها تغییرات غیر خطی دارند و این مدل پیش‌بینی قابل قبولی را ارائه نمی‌دهد.



شکل ۴-۲۸ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - سدکارده

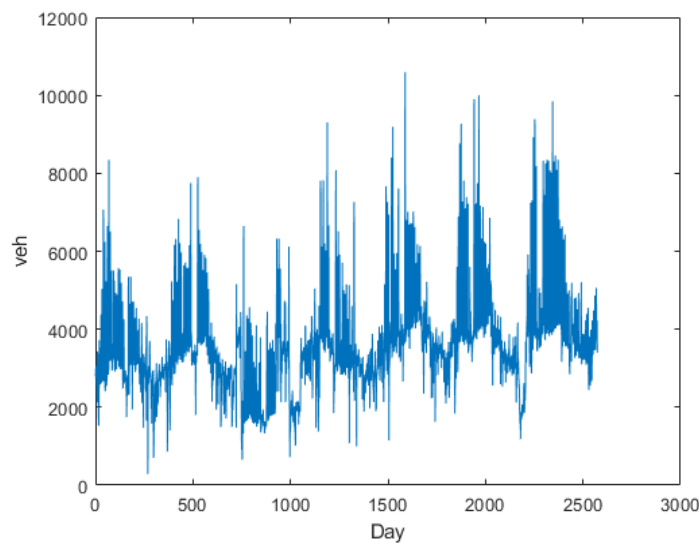
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۲۴ آمده است.

جدول ۴-۲۴ مقادیر خطا در محور مشهد- کلات

خطا	شاخص
MSE	۰,۴۰۸
RMSE	۰,۶۳۹
MAPE%	۱۶,۱

مسیر کلات - مشهد

شکل ۴-۲۹ حجم تردد روزانه برای ۲۹۵۴ روز در محور کلات - مشهد مشخص شده است.



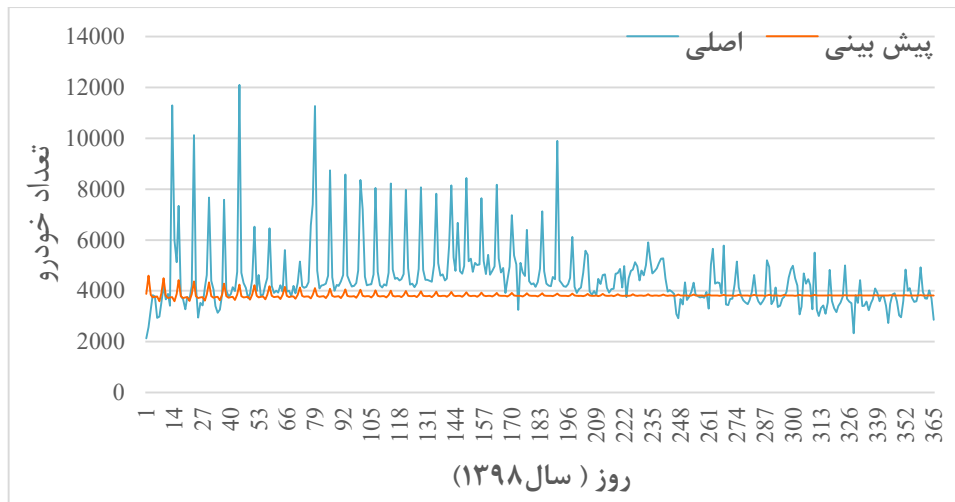
شکل ۴-۲۹ حجم تردد روزانه مسیر کلات - مشهد

برای داده‌های حجم تردد روزانه جاده‌ای در محور کلات - مشهد مدل به صورت $ARIMA(3,1,2)(1,0,1)$ به دست آمده است (جدول ۴-۲۵).

جدول ۴-۲۵ پارامترهای مدل برای محور کلات - مشهد با حذف ضریب ثابت

پارامترها	تاخیر	مولفه‌ها	تخمین	SE	t	Sig.
AR	Lag ۱	p=۱	۱,۰۴۲	۰,۰۲۰	۵۱,۵۶۴	۰,۰۰۰
	Lag ۲	p=۲	-۰,۲۴۰	۰,۰۲۸	-۸,۴۸۰	۰,۰۰۰
	Lag ۳	p=۳	۰,۰۶۱	۰,۰۲۰	۳,۰۲۶	۰,۰۰۳
Difference		d=۱	۱			
MA	Lag ۱	q=۱	۱,۶۸۱	۰,۰۰۴	۴۳۹,۰۷۵	۰,۰۰۰
	Lag ۲	q=۲	-۰,۶۸۱	۰,۰۰۳	-۲۰۲,۰۹۰	۰,۰۰۰
AR, Seasonal	Lag ۱	P=۱	۰,۹۲۰	۰,۰۱۲	۷۵,۸۲۸	۰,۰۰۰
MA, Seasonal	Lag ۱	Q=۱	۰,۶۴۹	۰,۰۲۳	۲۷,۶۱۱	۰,۰۰۰

در جدول ۴-۲۵ با توجه به خروجی جدول، چون مقدار معناداری یعنی Sig. کوچکتر از سطح آزمون، یعنی ۰/۰۵ شده است، فرضیه صفر را رد و فرض صفر بودن مقدار ضریب ثابت مدل تایید می‌شود. شکل ۴-۳۰ نمودار تغییرات حجم تردد روزانه نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد.



شکل ۴-۳۰ نمودار تغییرات حجم تردد نسبت به زمان در محور کلات - مشهد

همانطور که از شکل ۴-۳۰ مشخص می‌باشد به دلیل کاربری خاص این مسیر (توریستی - تفریحی) داده‌های اصلی در روزهای غیر تعطیل از یک روند خاص و خطی پیروی می‌کنند که روش ARIMA این روند را تشخیص داده است و پیش‌بینی قابل قبولی را ارائه می‌دهد اما برای داده‌های آخر هفته و مناسب‌هایی که تعطیلی رسمی هستند داده‌ها تغییرات غیر خطی دارند و این مدل پیش‌بینی قابل قبولی را ارائه نمی‌دهد. مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۲۶ آمده است.

جدول ۴-۲۶ مقادیر خطا در محور کلات - مشهد

خطا	شاخص
MSE	۰,۰۵۹
RMSE	۰,۲۴۳
MAPE%	۱۶,۵

۴ - ۵ مدل سازی و پیش بینی شبکه عصبی مصنوعی

همانطور که در فصل قبل بیان گردید، اولین قدم فراخوانی داده ها می باشد. در داده های اصلی حجم تردد، از روز اول تا یکسال پایانی به عنوان داده های ورودی و داده های یکسال پایانی به عنوان داده های هدف برای مدل تعریف می شوند. بنابراین شبکه عصبی مصنوعی مورد نظر برای پیش بینی حجم تردد یکسال پایانی آموزش می بیند. پس از انجام مدل سازی جداول و تصاویر مربوط به پیش بینی از قبیل جداول مربوط به پیش بینی یکسال پایانی، نمودار داده ها نسبت به زمان، نمودار رگرسیون خطی برای داده های آموزش و آزمایش، نمودار حجم تردد روزانه سال ۹۸، نمودار پیش بینی حجم تردد روزانه سال ۹۸ و نمودار خطا پیش بینی، توسط نرم افزار ارائه می شود.

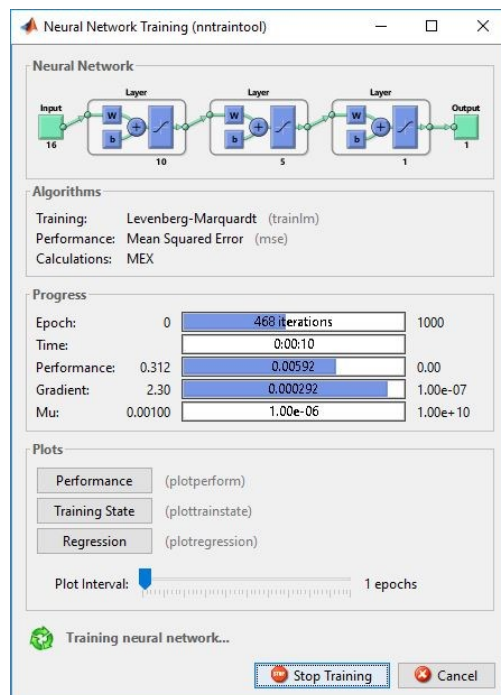
در جدول ۴-۲۷ زمان آموزش شبکه عصبی مصنوعی در نرم افزار Matlab ارائه شده است.

جدول ۴-۲۷ زمان مدل سازی در مدل شبکه عصبی مصنوعی

نام مسیر	زمان مدل سازی (s)	نام مسیر	زمان مدل سازی (s)
مشهد - سرخس	۱۸	سرخس - مشهد	۱۷
مشهد - چناران	۱۸	چناران - مشهد	۲۰
مشهد - فریمان	۱۷	فریمان - مشهد	۱۷
مشهد - ملک آباد	۱۹	ملک آباد - مشهد	۱۸
مشهد - سد کارده	۱۶	سد کارده - مشهد	۱۶
مشهد - کلات	۱۵	کلات - مشهد	۱۴

شکل ۴-۳۱ شرط های توقف شبکه را ارائه می دهد. در این تولباکس بدلیل اینکه نمی توان آموزش را تا ابد ادامه داد، مواردی مشخص می شود که اگر هر کدام از آن ها عملی شود اجرای آموزش متوقف می شود:

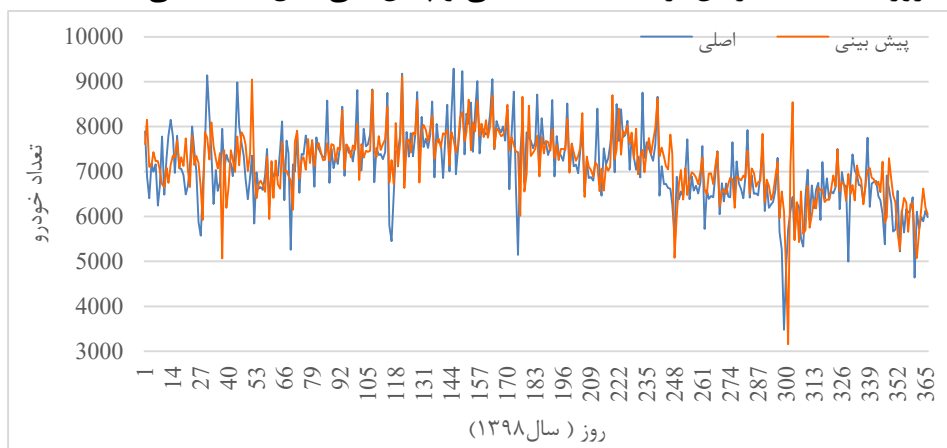
- Epoch (تکرار) : مثلاً بعد از ۱۰۰ بار تکرار متوقف شود
- Time (زمان) : محدودیت زمانی داشته باشیم
- Performance (عملکرد شبکه) : اگر بتوانیم میزان خطا و اختلاف بین نتیجه ای که از شبکه عصبی می گیریم با نتیجه واقعی از یک مقداری کمتر کنیم اجرا متوقف می شود.
- Gradient (گرادیان) : میزان مشتق موجود در شبکه عصبی
- Mu : شاخص های الگوریتم بهینه سازی که مورد استفاده قرار گرفته است.



شکل ۴-۳۱ نمونه‌ای از اجرای شبکه عصبی مصنوعی

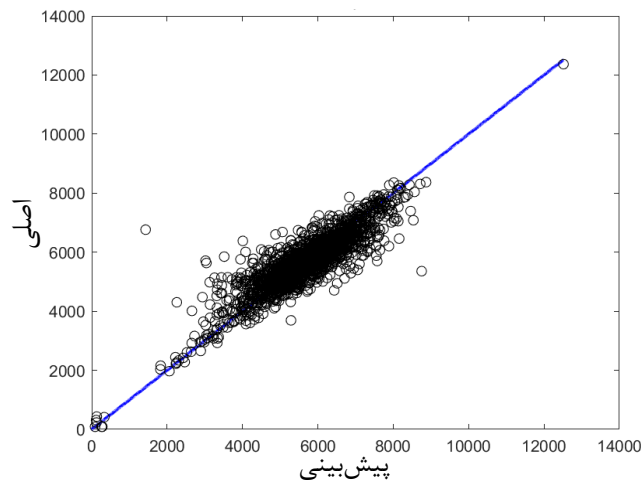
مسیر مشهد - سرخس

در جدول ۴-۱ بازه زمانی برداشت داده برای هر مسیر بیان گردیده است و همچنین نمودار داده‌های حجم تردد نسبت به زمان برای ۳۳۱۶ روز در شکل ۴-۷ ترسیم شده است. شکل ۴-۳۲ نمودار تغییرات حجم تردد روزانه نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد.



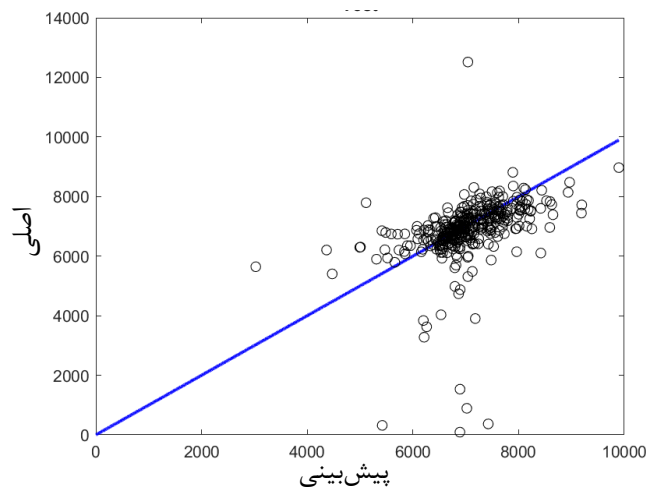
شکل ۴-۳۲ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - سرخس

در شکل ۴-۳۲ خطوط آبی، مقادیر حجم تردد برداشت شده و خطوط نارنجی، مقادیر حجم تردد پیش‌بینی شده در سال ۱۳۹۸ در محور مورد نظر می‌باشد. همچنین محور عمودی برای نمودار پیش-بینی حجم تردد روزانه می‌باشد و محور افقی بیانگر روزهای سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، اگرچه در نگاه اول پیش‌بینی خوبی مشاهده می‌شود ولی مقادیر حجم تردد پیش‌بینی شده در این روش دارای جهش‌های (افزایشی یا کاهش) غیر قابل چشم‌پوشی می‌باشد.



شکل ۴-۳۳ نمودارهای رگرسیون خطی برای داده های آموزش

در شکل ۴-۳۳ و ۴-۳۴ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. محور عمودی این نمودارها حجم تردد برداشت شده و محور افقی حجم تردد پیش‌بینی شده توسط مدل می‌باشد. همانطور که در این شکل‌های مشخص می‌باشد، داده‌های آزمایشی برای تقریباً ۳۰۰۰ روز از رگرسیون خطی خوبی برخوردارند اما داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی نسبت به داده‌های آزمایشی برخوردار نیست.



شکل ۴-۳۴ نمودارهای رگرسیون خطی برای داده های آزمایش

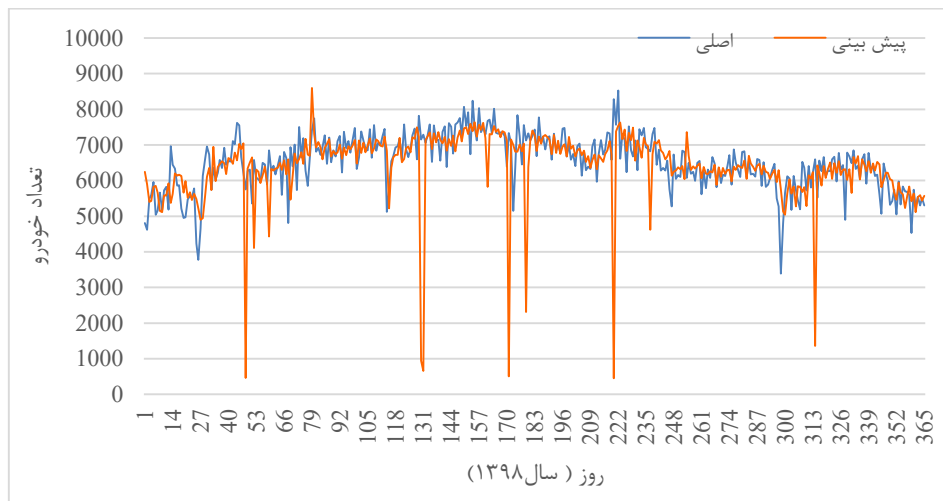
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۲۸ آمده است.

جدول ۴-۲۸ مقادیر خطا در محور مشاهد - سرخس

خطا	شاخص
۰,۰۲۸	MSE
۰,۱۶۶	RMSE
۶,۵	MAPE%

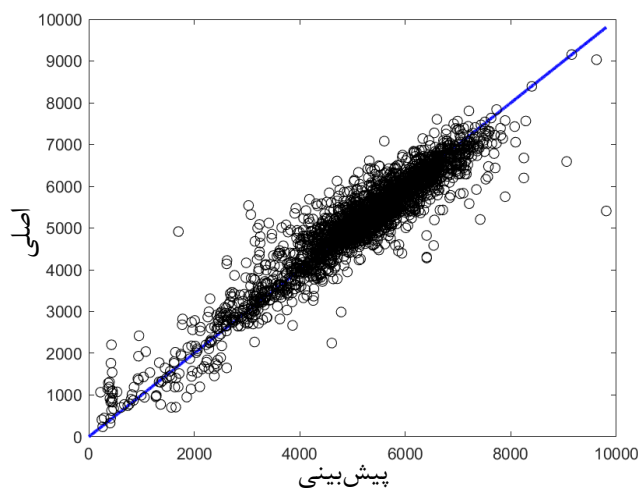
مسیر سرخس - مشهد

نمودار داده‌های حجم تردد نسبت به زمان برای ۳۳۱۶ روز در شکل ۴-۹ ترسیم شده است. در شکل ۴-۳۵ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد.



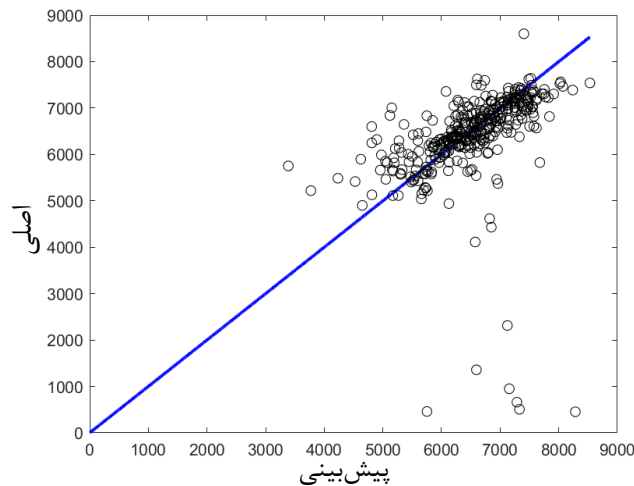
شکل ۴-۳۵ نمودار تغییرات حجم تردد نسبت به زمان در محور سرخس - مشهد

همانطور که در شکل ۴-۳۵ مشخص می‌باشد، اگرچه در نگاه اول پیش‌بینی خوبی مشاهده می‌شود ولی مقادیر حجم تردد پیش‌بینی شده در این روش دارای جهش‌های کاهشی غیر قابل چشم‌پوشی می‌باشد.



شکل ۴-۳۶ نمودارهای رگرسیون خطی برای داده‌های آموزش

در شکل ۴-۳۶ و ۴-۳۷ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در شکل‌ها مشخص می‌باشد داده‌های آزمایشی برای تقریباً ۳۰۰۰ روز از رگرسیون خطی خوبی برخوردارند اما داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی نسبت به داده‌های آزمایشی برخوردار نیست.



شکل ۴-۳۷ نمودارهای رگرسیون خطی برای داده های آزمایش

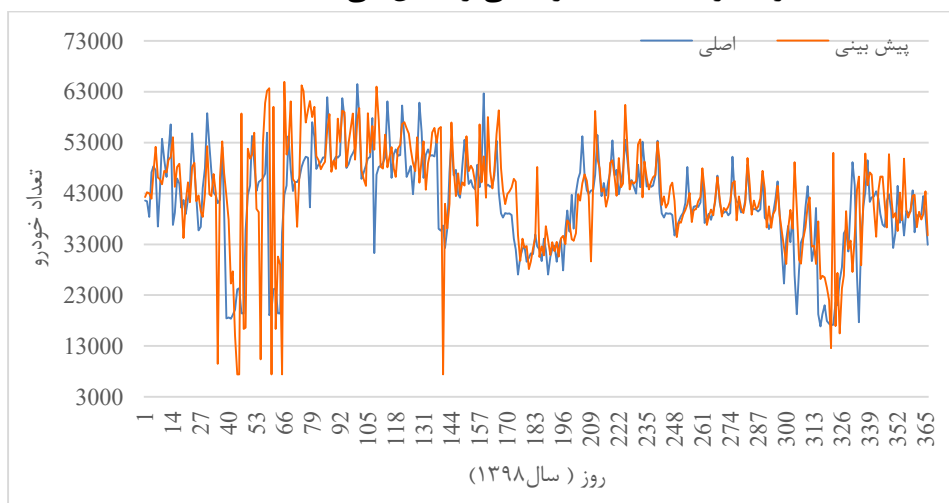
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۲۹ آمده است.

جدول ۴-۲۹ مقادیر خطا در محور سرخس - مشهد

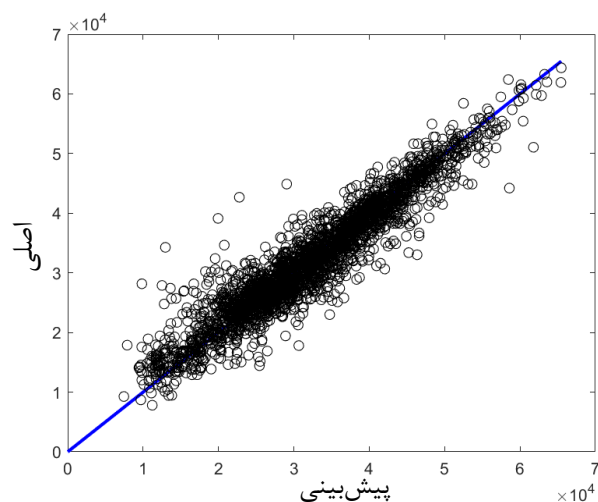
شاخص	خطا
MSE	۰,۰۴۵
RMSE	۰,۲۱۳
MAPE%	۷,۸

مسیر مشهد - چناران

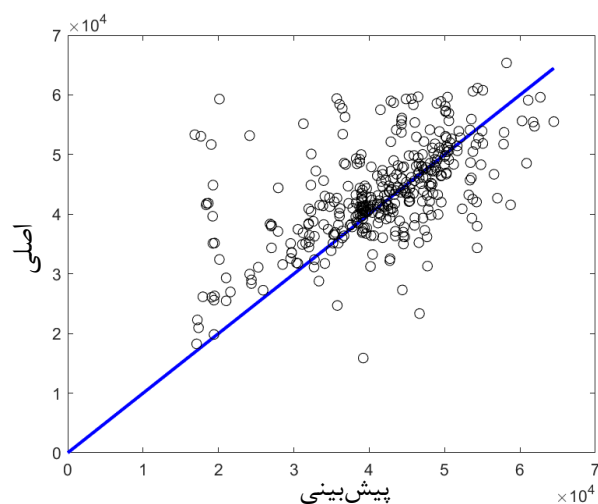
نمودار داده‌های حجم تردد روزانه نسبت به زمان برای ۳۶۵۲ روز در شکل ۴-۱۱ ترسیم شده است. در شکل ۴-۳۸ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، اگرچه این مدل وابستگی غیرخطی داده‌ها را تشخیص داده است اما خطای اندازه گیری شده مقادیر بالایی را نشان می‌دهند.



شکل ۴-۳۸ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - چناران



شکل ۴-۳۹ نمودارهای رگرسیون خطی برای داده های آموزش



شکل ۴-۴۰ نمودارهای رگرسیون خطی برای داده های آموزش

در شکل ۴-۳۹ و ۴-۴۰ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در اشکال مشخص می‌باشد داده‌های آزمایشی برای ۳۳۰۰ روز از رگرسیون خطی خوبی برخوردارند اما داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی نسبت به داده‌های آزمایشی برخوردار نیست.

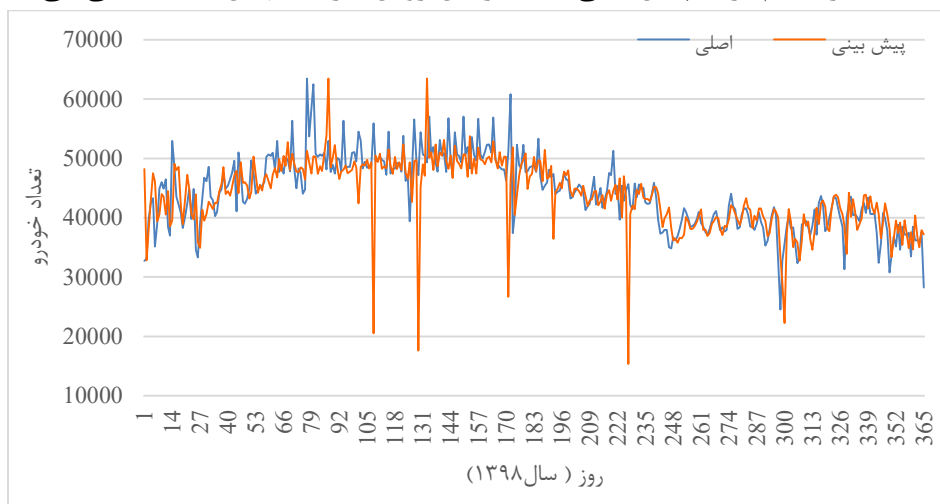
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۳۰ آمده است.

جدول ۴-۳۰ مقادیر خطا در محور مشهد - چناران

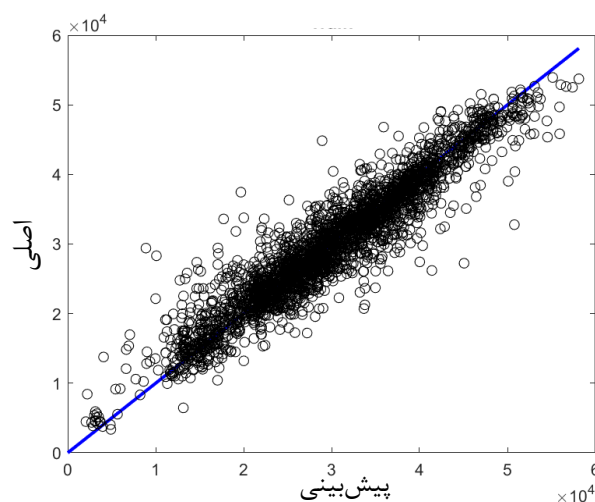
خطا	شاخص
۰,۰۸۵	MSE
۰,۲۹۲	RMSE
۱۶,۶	MAPE%

مسیر چناران - مشهد

نمودار داده‌های حجم تردد روزانه نسبت به زمان برای ۳۶۵۲ روز در شکل ۴-۱۳ ترسیم شده است. در شکل ۴-۴۱ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، اگرچه این مدل وابستگی غیرخطی داده‌ها را تشخیص داده است اما مقادیر حجم تردد پیش‌بینی شده در این روش دارای جهش‌های کاهشی می‌باشند.

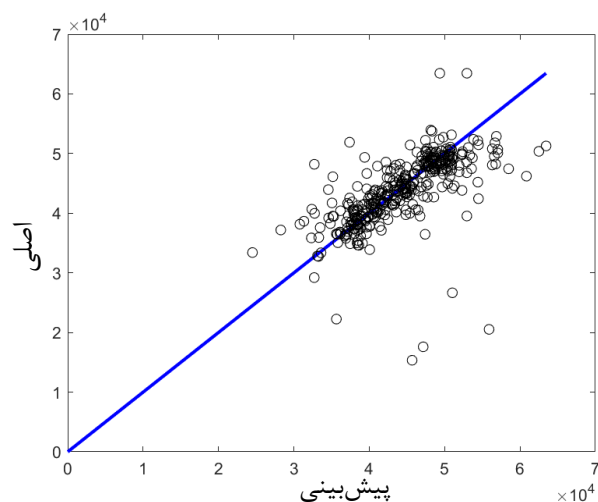


شکل ۴-۴۱ نمودار تغییرات حجم تردد نسبت به زمان در محور چناران - مشهد



شکل ۴-۴۲ نمودارهای رگرسیون خطی برای داده‌های آموزش

در شکل ۴-۴۲ و ۴۳ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در اشکال مشخص می‌باشد داده‌های آزمایشی برای ۳۳۰۰ روز از رگرسیون خطی خوبی برخوردارند اما داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی نسبت به داده‌های آزمایشی برخوردار نیست.



شکل ۴-۳ نمودارهای رگرسیون خطی برای داده های آزمایش

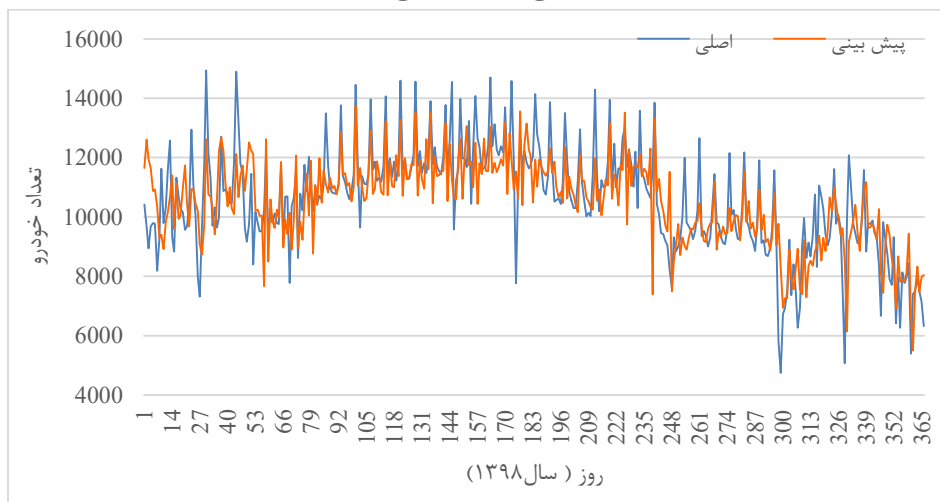
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۳۱ آمده است.

جدول ۴-۳۱ مقادیر خطا در محور چنان‌ان - مشهد

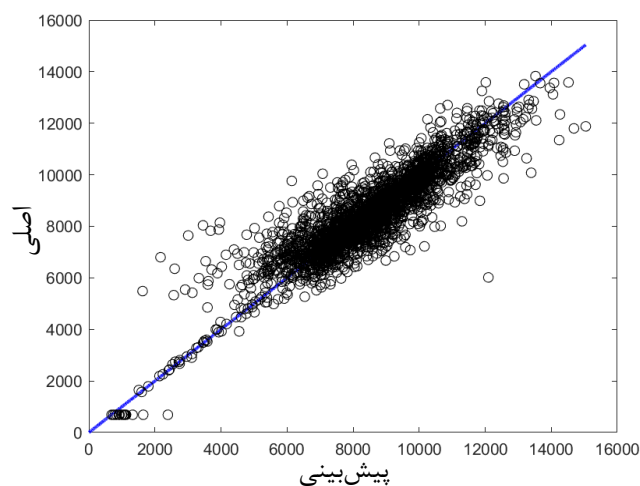
خطا	روش
۰,۰۲۶	MSE
۰,۱۶۱	RMSE
۶,۵	MAPE%

مسیر مشهد - فریمان

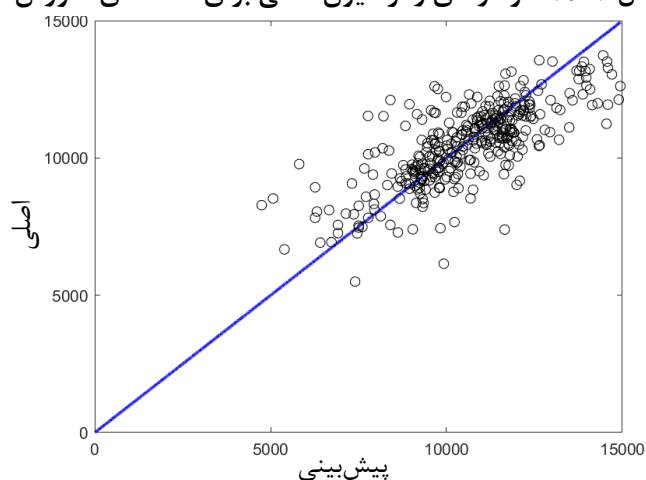
نمودار داده‌های حجم تردد نسبت به زمان برای ۳۳۱۶ روز در شکل ۴-۱۵ ترسیم شده است. در شکل ۴-۴۴ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، اگرچه این مدل وابستگی غیرخطی داده‌ها را تشخیص داده است. اما خطای اندازه‌گیری شده مقادیر بالایی را نشان می‌دهند.



شکل ۴-۴۴ نمودار تغییرات حجم تردد نسبت به زمان، پیش‌بینی و نمودار خطا در محور مشهد - فریمان



شکل ۴-۴۵ نمودارهای رگرسیون خطی برای داده های آموزش



شکل ۴-۴۶ نمودارهای رگرسیون خطی برای داده های آزمایش

در شکل ۴-۴۵ و ۴-۴۶ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در شکل مشخص می‌باشد داده‌های آزمایشی برای تقریباً ۳۰۰۰ روز از رگرسیون خطی خوبی برخوردارند اما داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی نسبت به داده‌های آزمایشی برخوردار نیست.

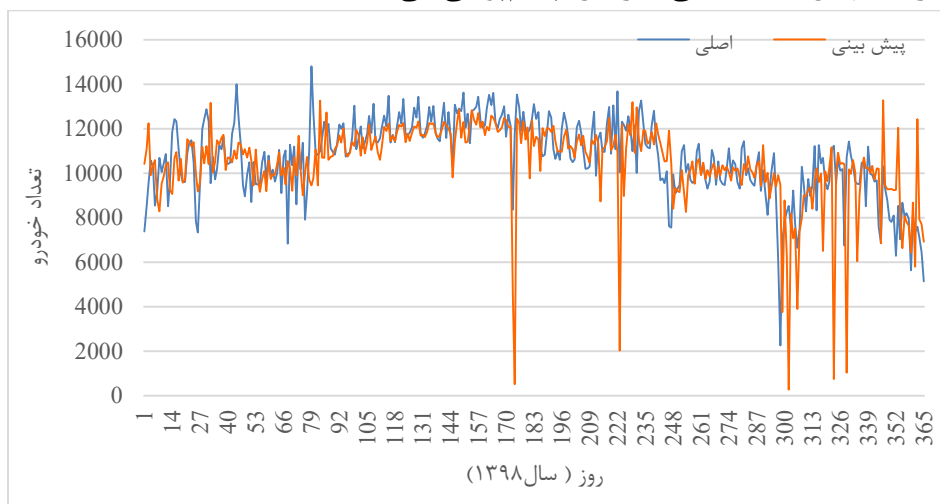
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۳۲ آمده است.

جدول ۴-۳۲ مقادیر خطا در محور مشهد – فریمان

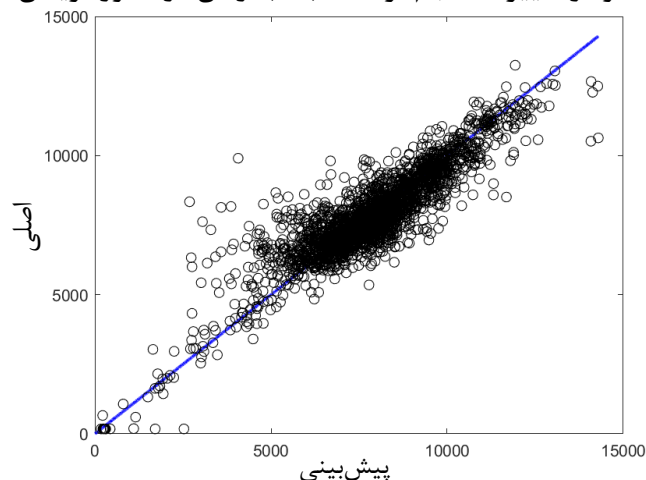
خطا	شاخص
۰,۰۲۷	MSE
۰,۱۶۶	RMSE
۸,۷	MAPE%

مسیر فریمان - مشهد

نمودار داده‌های حجم تردد نسبت به زمان برای ۳۳۱۶ روز در شکل ۴-۱۷ ترسیم شده است. در شکل ۴-۴۷ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، اگرچه این مدل وابستگی غیرخطی داده‌ها را تشخیص داده است. اما دارای جهش‌های کاهشی غیرقابل چشم‌پوشی می‌باشد.

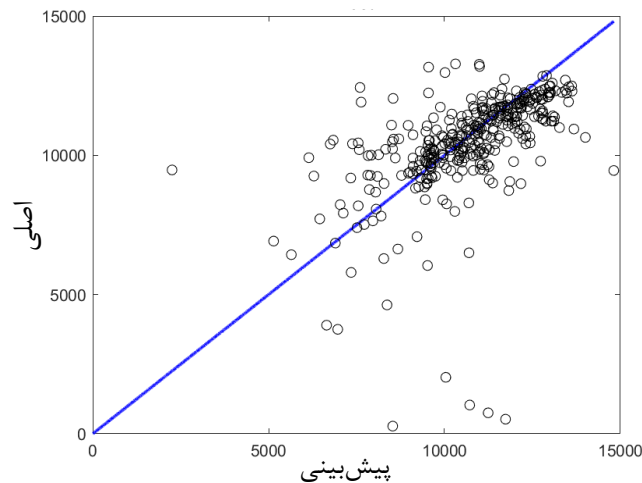


شکل ۴-۴۷ نمودار تغییرات حجم تردد نسبت به زمان در محور فریمان - مشهد



شکل ۴-۴۸ نمودارهای رگرسیون خطی برای داده‌های آموزش

در شکل ۴-۴۸ و ۴-۴۹ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در اشکال مشخص می‌باشد داده‌های آزمایشی برای ۳۰۰۰ روز از رگرسیون خطی خوبی برخوردارند اما داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی نسبت به داده‌های آزمایشی برخوردار نیست.



شکل ۴-۴۹ نمودارهای رگرسیون خطی برای داده های آزمایش

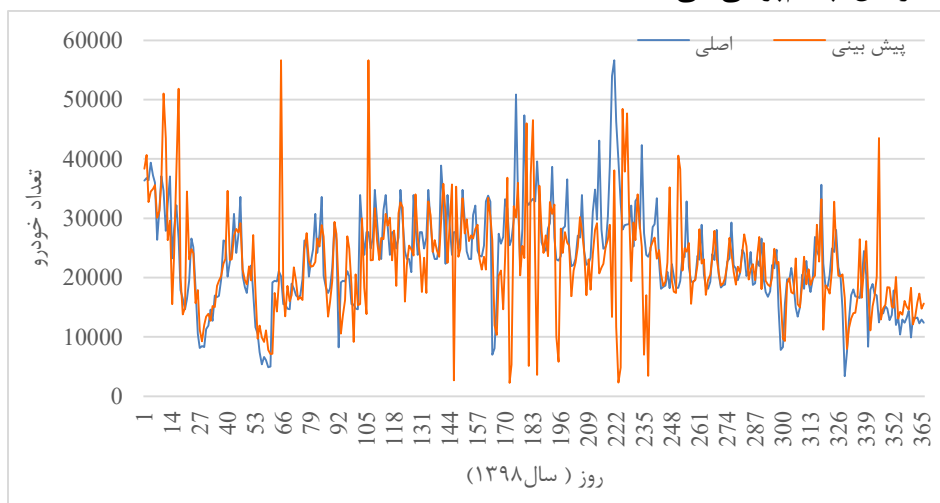
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۳۳ آمده است.

جدول ۴-۳۳ مقادیر خطا در محور فریمان - مشهد

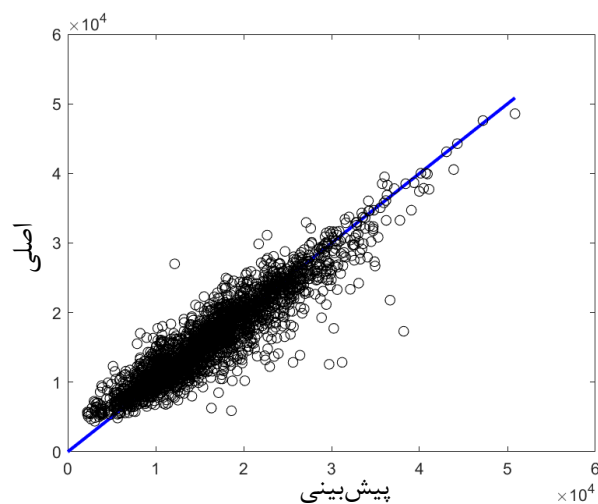
شاخص	خطا
MSE	۰,۰۵۷
RMSE	۰,۲۳۹
MAPE%	۱۱,۲

مسیر مشهد - ملک آباد

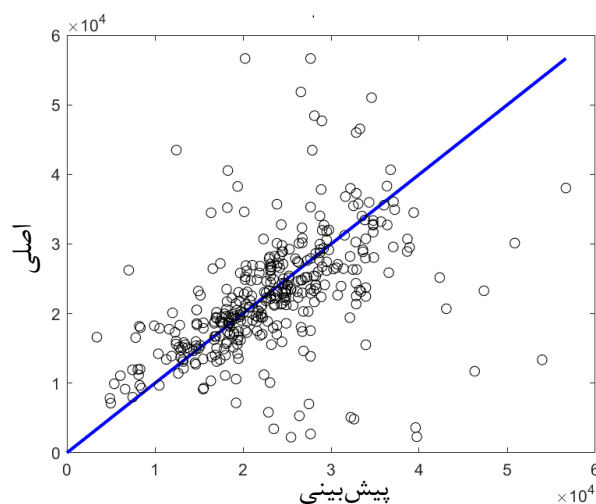
نمودار داده‌های حجم تردد نسبت به زمان برای ۳۶۵۲ روز در شکل ۴-۱۹ ترسیم شده است. در شکل ۴-۵۰ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، اگرچه این مدل وابستگی غیرخطی داده‌ها را تشخیص داده است. اما خطای اندازه‌گیری شده مقادیر بالایی را نشان می‌دهند. همچنین دارای جهش‌های افزایشی و کاهشی غیرقابل چشم‌پوشی می‌باشد.



شکل ۴-۵۰ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - ملک آباد



شکل ۴-۵۱ نمودارهای رگرسیون خطی برای داده های آموزش



شکل ۴-۵۲ نمودارهای رگرسیون خطی برای داده های آزمایش

در شکل ۴-۵۱ و ۴-۵۲ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در اشکال مشخص می‌باشد داده‌های آزمایشی برای ۳۳۰۰ روز از رگرسیون خطی خوبی برخوردارند اما داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی نسبت به داده‌های آزمایشی برخوردار نیست.

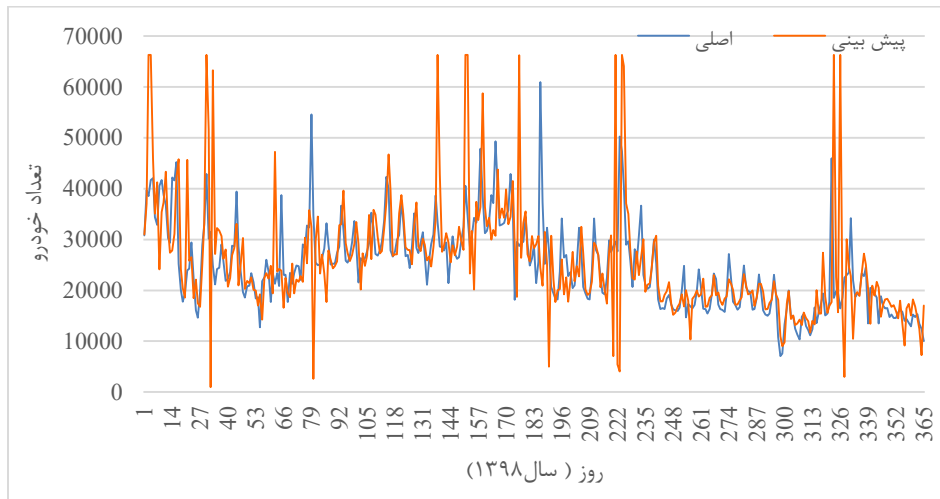
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۳۴ آمده است.

جدول ۴-۳۴ مقادیر خطا در محور مشهد - ملک‌آباد

خطا	شاخص
MSE	۰,۰۹۹
RMSE	۰,۳۱۴
MAPE%	۲۵,۱

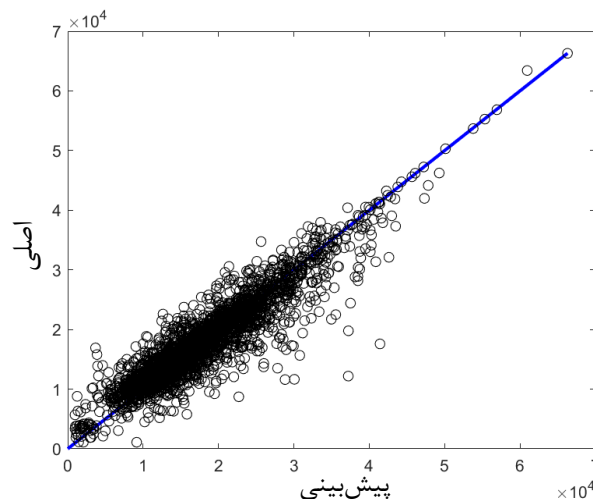
مسیر ملک آباد - مشهد

نمودار داده‌های حجم تردد نسبت به زمان برای ۳۶۵۲ روز در شکل ۴-۲۱ ترسیم شده است. در شکل ۴-۵۳ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد.



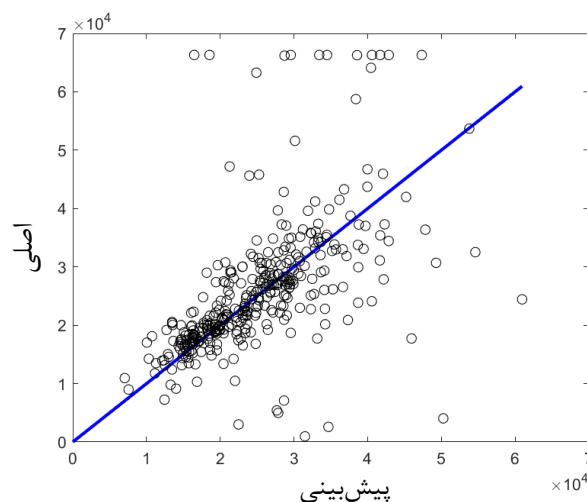
شکل ۴-۵۳ نمودار تغییرات حجم تردد نسبت به زمان در محور ملک آباد - مشهد

همانطور که در شکل ۴-۵۳ مشخص می‌باشد، اگرچه این مدل وابستگی غیرخطی داده‌ها را تشخیص داده است. اما خطای اندازه‌گیری شده مقادیر بالایی را نشان می‌دهند. همچنین دارای جهش‌های افزایشی و کاهش‌ی غیرقابل چشم‌پوشی می‌باشد.



شکل ۴-۵۴ نمودارهای رگرسیون خطی برای داده‌های آموزش

در شکل ۴-۵۴ و ۴-۵۵ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در اشکال مشخص می‌باشد داده‌های آزمایشی برای ۳۳۰۰ روز از رگرسیون خطی خوبی برخوردارند اما داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی نسبت به داده‌های آزمایشی برخوردار نیست.



شکل ۴-۵۵ نمودارهای رگرسیون خطی برای داده‌های آزمایش

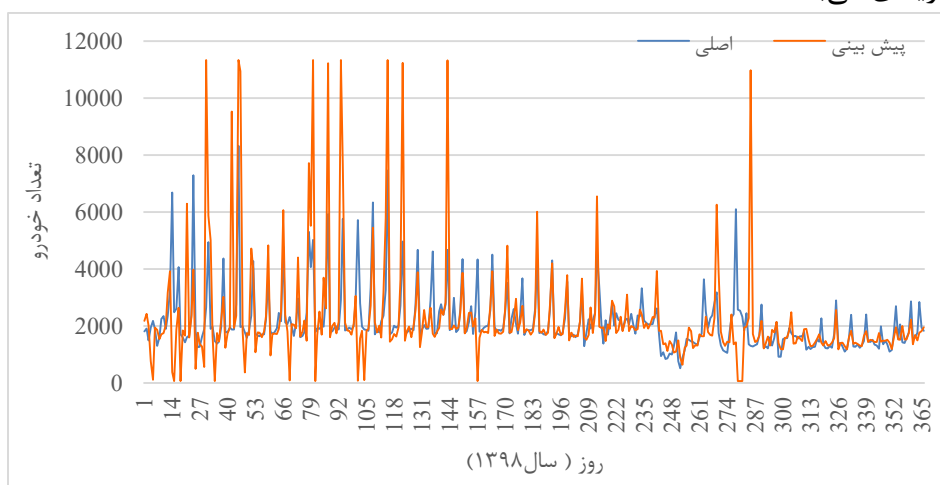
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۳۵ آمده است.

جدول ۴-۳۵ مقادیر خطا در محور ملک‌آباد - مشهد

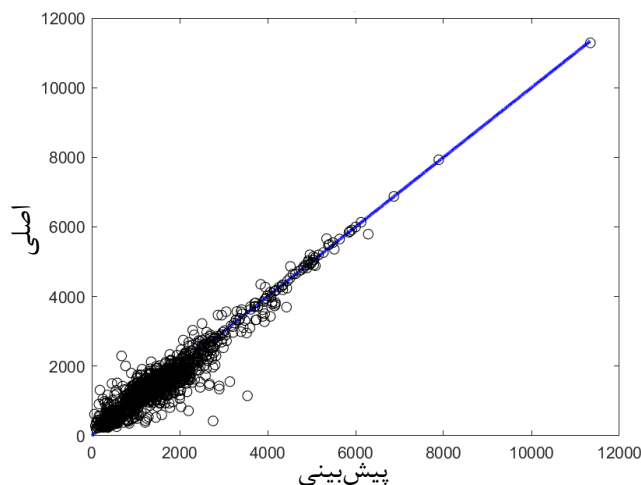
خطا	شاخص
MSE	۰,۰۸۵
RMSE	۰,۲۹۲
MAPE%	۲۰,۶

مسیر مشهد - سدکارده

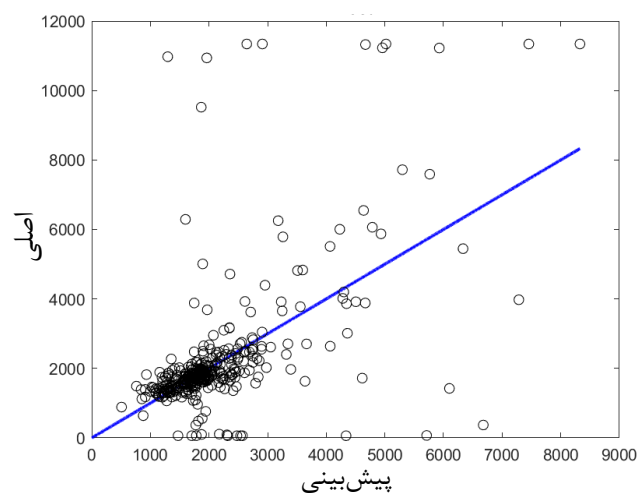
نمودار داده‌های حجم تردد نسبت به زمان برای ۳۱۳۲ روز در شکل ۴-۲۳ ترسیم شده است. در شکل ۴-۵۶ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، اگرچه این مدل وابستگی غیرخطی داده‌ها را تشخیص داده است. اما خطای اندازه‌گیری شده مقادیر بالایی را نشان می‌دهند. همچنین دارای جهش‌های افزایشی و کاهشی زیادی می‌باشد.



شکل ۴-۵۶ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - سدکارده



شکل ۴-۵۷ نمودارهای رگرسیون خطی برای داده های آموزش



شکل ۴-۵۸ نمودارهای رگرسیون خطی برای داده های آزمایش

در شکل ۴-۵۷ و ۴-۵۸ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در شکل مشخص می‌باشد داده‌های آزمایشی برای ۲۷۶۷ روز از رگرسیون خطی خوبی برخوردارند اما داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی نسبت به داده‌های آزمایشی برخوردار نیست.

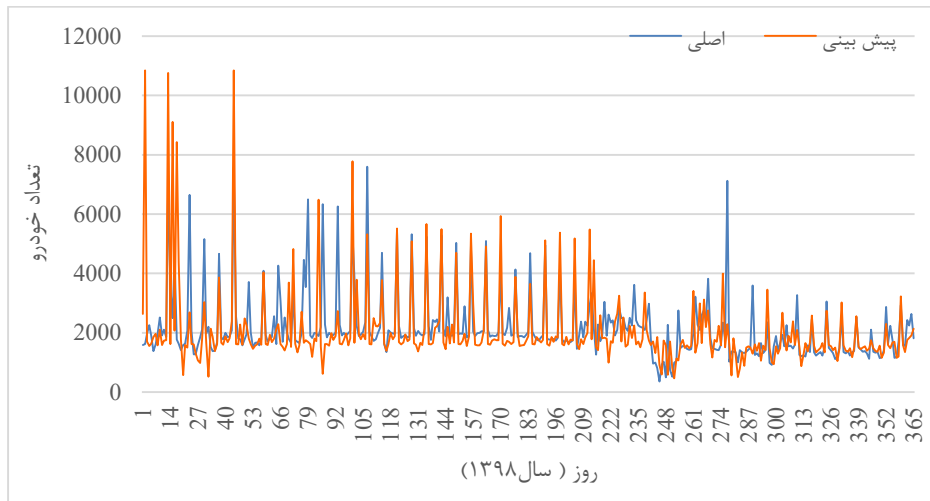
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۳۶ آمده است.

جدول ۴-۳۶ مقادیر خطا در محور مشهد - سدکارده

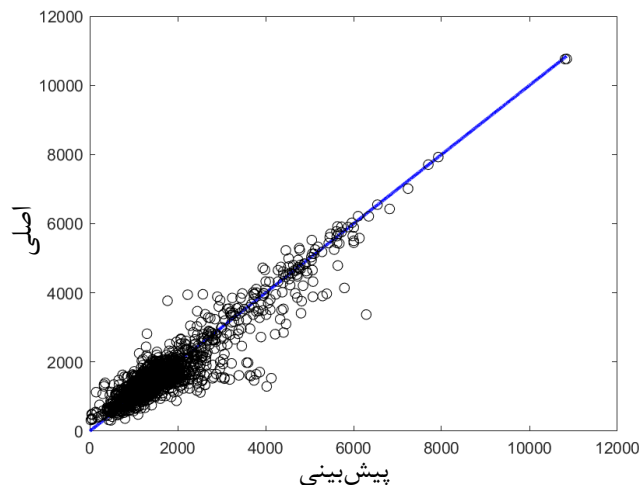
خطا	شاخص
۰,۰۷۵	MSE
۰,۲۷۵	RMSE
۲۹,۳	MAPE%

مسیر سدکارده - مشهد

نمودار داده‌های حجم تردد نسبت به زمان برای ۳۱۳۲ روز در شکل ۴-۲۵ ترسیم شده است. در شکل ۴-۵۹ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، اگرچه این مدل وابستگی غیرخطی داده‌ها را تشخیص داده است. اما خطای اندازه‌گیری شده مقادیر بالایی را نشان می‌دهند. همچنین دارای جهش‌های افزایشی و کاهشی زیادی می‌باشد.

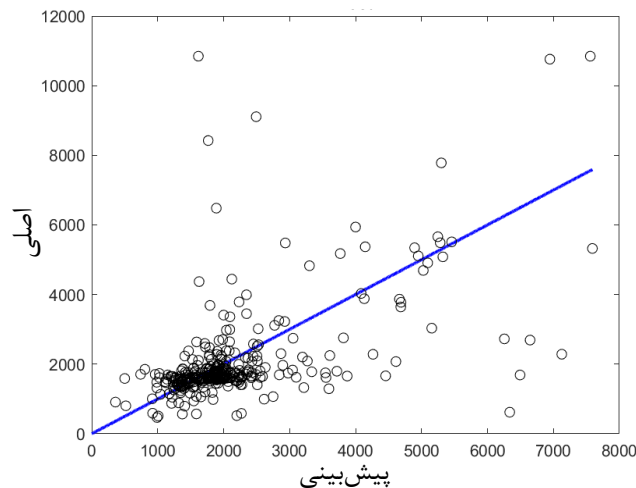


شکل ۴-۵۹ نمودار تغییرات حجم تردد نسبت به زمان در محور سدکارده - مشهد



شکل ۴-۶۰ نمودارهای رگرسیون خطی برای داده‌های آموزش

در شکل ۴-۶۰ و ۴-۶۱ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در اشکال مشخص می‌باشد داده‌های آزمایشی برای ۲۷۶۷ روز از رگرسیون خطی خوبی برخوردارند اما داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی نسبت به داده‌های آزمایشی برخوردار نیست.



شکل ۴-۶۱ نمودارهای رگرسیون خطی برای داده‌های آزمایش

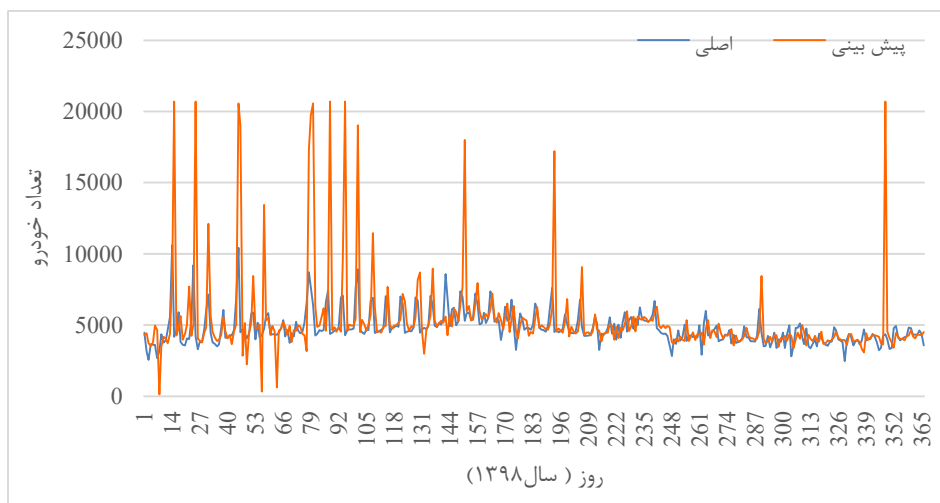
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۳۷ آمده است.

جدول ۴-۳۷ مقادیر خطا در محور سداکارده - مشهد

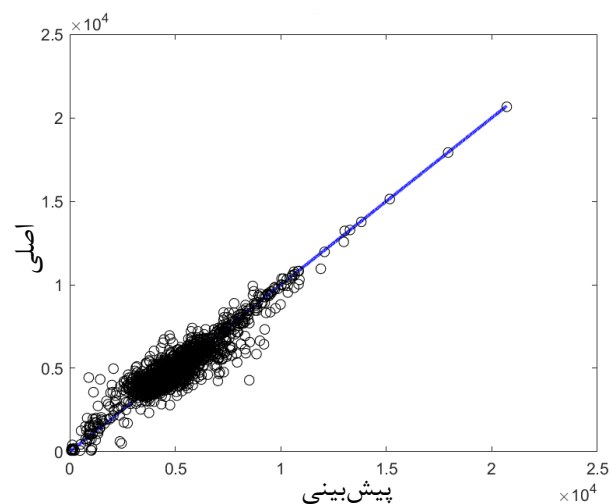
شاخص	خطا
MSE	۰,۰۴۷
RMSE	۰,۲۱۸
MAPE%	۲۷,۸

مسیر مشهد - کلات

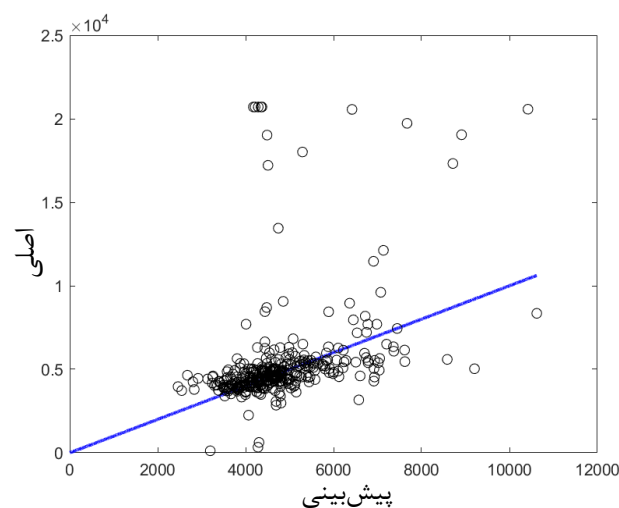
نمودار داده‌های حجم تردد نسبت به زمان برای ۲۹۵۴ روز در شکل ۴-۲۷ ترسیم شده است. در شکل ۴-۶۲ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال می‌باشد. همانطور که در شکل مشخص می‌باشد، اگرچه این مدل وابستگی غیرخطی داده‌ها را تشخیص داده است. اما خطای اندازه‌گیری شده مقادیر بالایی را نشان می‌دهند. همچنین دارای جهش‌های افزایشی و کاهشی زیادی می‌باشد.



شکل ۴-۶۲ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - کلات



شکل ۴-۶۳ نمودارهای رگرسیون خطی برای داده های آموزش



شکل ۴-۶۴ نمودارهای رگرسیون خطی برای داده های آزمایش

در شکل ۴-۶۳ و ۴-۶۴ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در اشکال مشخص می‌باشد داده‌های آزمایشی برای ۲۵۹۰ روز از رگرسیون خطی خوبی برخوردارند اما داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی نسبت به داده‌های آزمایشی برخوردار نیست.

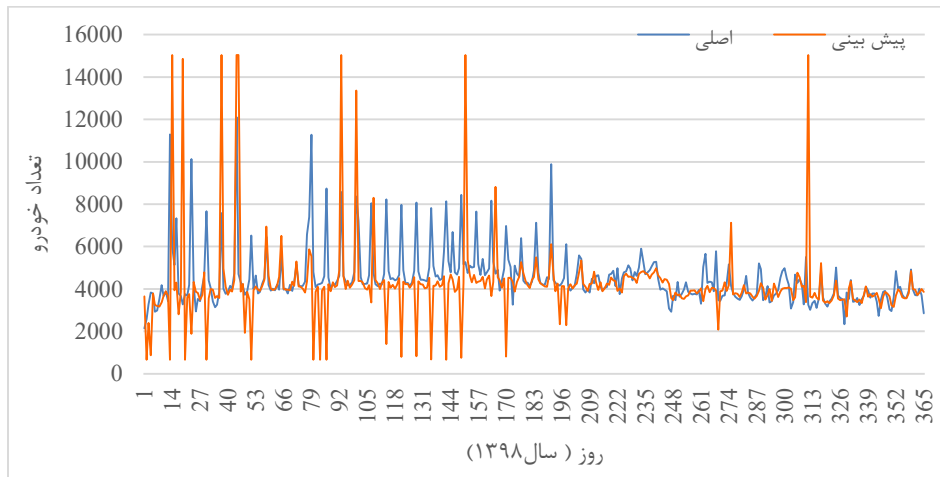
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۳۸ آمده است.

جدول ۴-۳۸ مقادیر خطا در محور مشهود - کلات

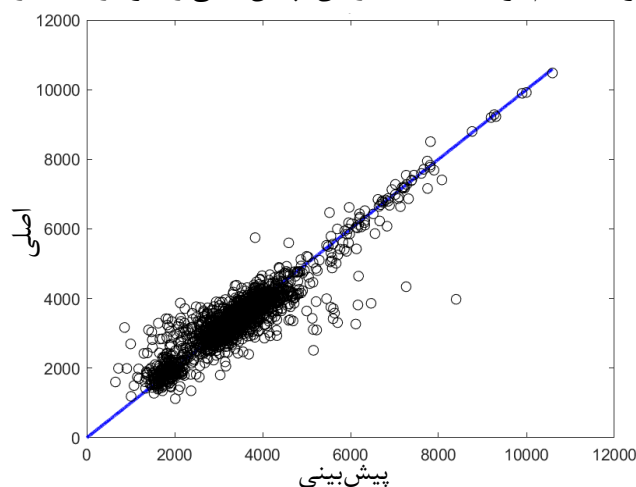
خطا	شاخص
۰,۰۷۷	MSE
۰,۲۷۷	RMSE
۲۴,۲	MAPE%

مسیر کلات - مشهد

نمودار داده‌های حجم تردد نسبت به زمان برای ۲۹۵۴ روز در شکل ۴-۲۹ ترسیم شده است. در شکل ۴-۶۵ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال می‌باشد. همانطور که در شکل مشخص می‌باشد، اگرچه این مدل وابستگی غیرخطی داده‌ها را تشخیص داده است. اما خطای اندازه‌گیری شده مقادیر بالایی را نشان می‌دهند. همچنین دارای جهش‌های افزایشی و کاهشی زیادی می‌باشد.

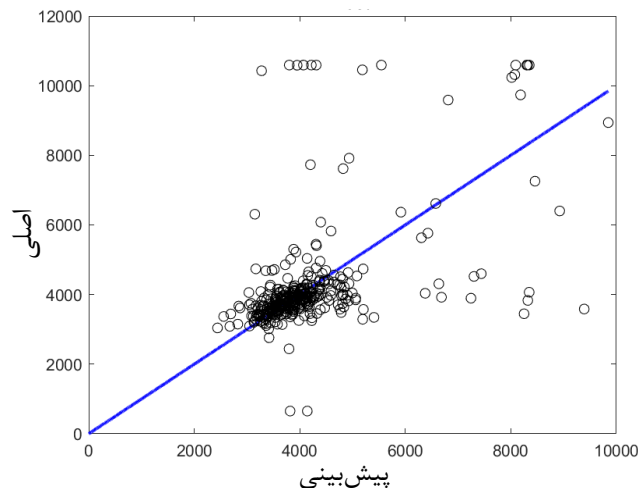


شکل ۴-۶۵ نمودار تغییرات حجم تردد نسبت به زمان، پیش‌بینی و نمودار خطا در محور کلات - مشهد



شکل ۴-۶۶ نمودارهای رگرسیون خطی برای داده‌های آزمایش

شکل ۴-۶۶ و ۴-۶۷ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در اشکال مشخص می‌باشد داده‌های آزمایشی برای ۲۵۹۰ روز از رگرسیون خطی خوبی برخوردارند اما داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی نسبت به داده‌های آزمایشی برخوردار نیست.



شکل ۴-۶۷ نمودارهای رگرسیون خطی برای داده های آزمایش

مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۳۹ آمده است.

جدول ۴-۳۹ مقادیر خطا در محور کلات - مشهد

خطا	روش
MSE	۰,۰۷۲
RMSE	۰,۲۶۹
MAPE%	۱۸,۳

۴ - ۶ مدل سازی و پیش‌بینی یادگیری عمیق LSTM

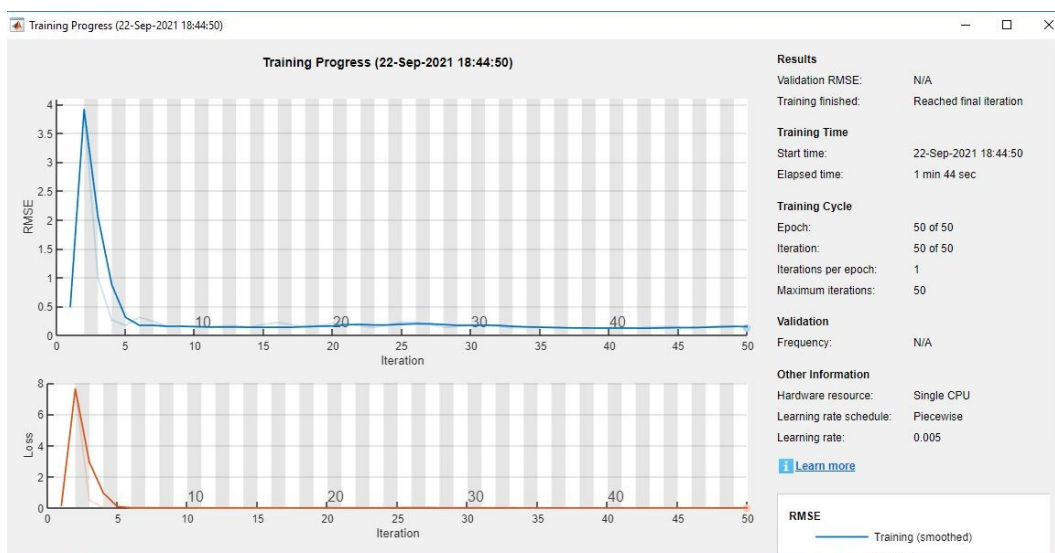
همانطور که در فصل قبل بیان گردید، تمام مراحل انجام شبکه عصبی مصنوعی و یادگیری عمیق LSTM به جز مرحله تشکیل ساختار شبکه، با یکدیگر مشابه کدنویسی شده‌اند. شبکه عمیق LSTM همانند شبکه عصبی مصنوعی برای پیش‌بینی یکسال پایانی حجم تردد روزانه مسیرهای مورد نظر آموزش می‌بیند. پس از انجام مدل سازی از آنجایی که داده‌ها برای شبکه عصبی مصنوعی و یادگیری عمیق LSTM یکسان می باشد، نمودار حجم تردد روزانه برای هر مسیر یکسان می باشد و از خروجی کد MATLAB شبکه عمیق LSTM حذف می‌شود. جداول و تصاویر مربوط به پیش‌بینی از قبیل جداول مربوط به داده‌های پیش‌بینی یکسال پایانی، نمودار رگرسیون خطی برای داده‌های آموزش و آزمایش، نمودار حجم تردد روزانه سال ۹۸، نمودار پیش‌بینی حجم تردد روزانه سال ۹۸ و نمودار خطا پیش‌بینی توسط نرم افزار ارائه می شود.

در جدول ۴-۴۰ زمان آموزش شبکه LSTM در نرم افزار Matlab ارائه شده است.

جدول ۴-۴۰ زمان مدلسازی در شبکه LSTM

نام مسیر	زمان مدلسازی (s)	نام مسیر	زمان مدلسازی (s)
مشهد - سرخس	۱۰۹	سرخس - مشهد	۱۱۱
مشهد - چناران	۱۳۱	چناران - مشهد	۱۲۳
مشهد - فریمان	۱۱۱	فریمان - مشهد	۱۱۳
مشهد - ملک آباد	۱۲۲	ملک آباد - مشهد	۱۳۳
مشهد - سد کارده	۱۰۶	سد کارده - مشهد	۱۰۴
مشهد - کلات	۱۰۶	کلات - مشهد	۸۴

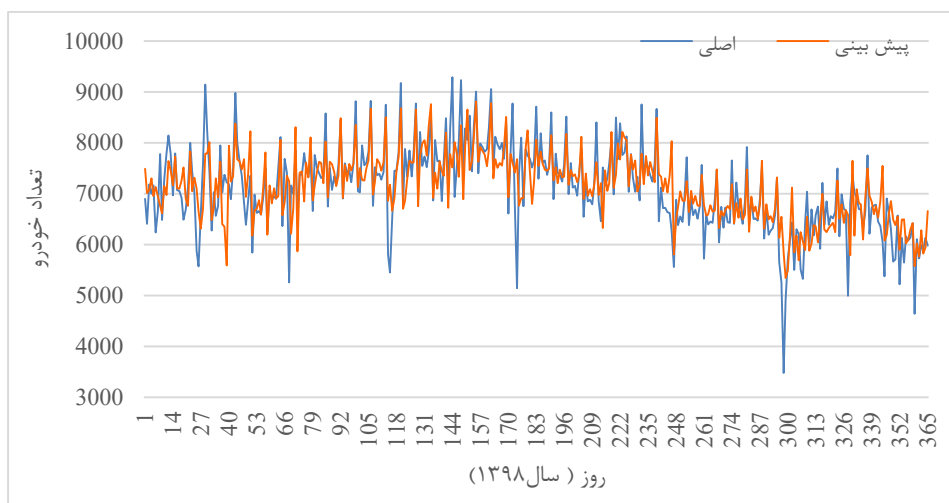
در شکل ۴-۶۸ شرط توقفی برای شبکه LSTM تعریف نشده است. در شبکه LSTM داده‌ها به ۵۰ دسته تقسیم شده‌اند و در ۵۰ مرتبه تکرار به شبکه ارائه می‌شود. در شکل Epoch بیانگر تعداد کل تکرار در داده‌های آموزش، Iteration بیانگر تعداد کل دسته‌ها می‌باشد.



شکل ۴-۶۸ نمونه‌ای از اجرای شبکه یادگیری عمیق LSTM

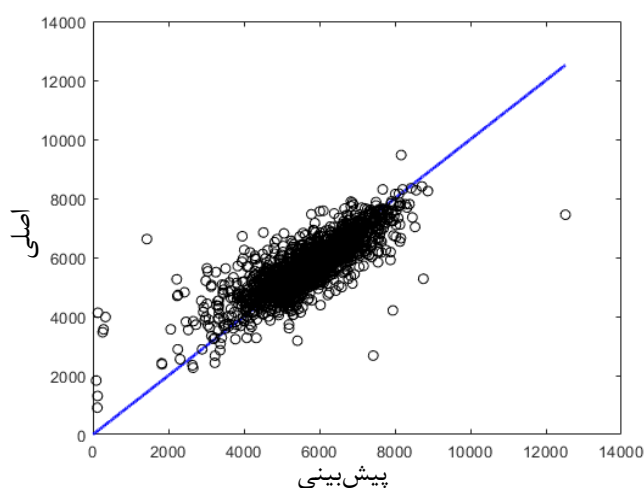
مسیر مشهد - سرخس

جدول ۴-۱ بازه زمانی برداشت داده برای هر مسیر بیان گردیده است و همچنین نمودار داده‌های حجم تردد نسبت به زمان برای ۳۳۱۶ روز در شکل ۴-۷ ترسیم شده است. در شکل ۴-۶۹ نمودار تغییرات حجم تردد روزانه نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. در شکل خطوط آبی، مقادیر حجم تردد برداشت شده و خطوط قرمز، مقادیر حجم تردد پیش‌بینی شده در سال ۱۳۹۸ در محور مورد نظر می‌باشد. همچنین محور عمودی حجم تردد روزانه می‌باشد و محور افقی بیانگر روزهای سال ۱۳۹۸ می‌باشد.

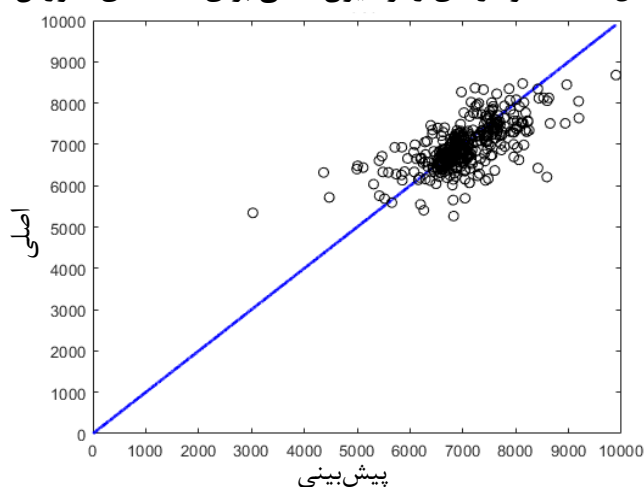


شکل ۴-۶۹ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - سرخس

همانطور که در شکل ۴-۶۹ مشخص می باشد، مدل وابستگی های فصلی و غیر خطی را تشخیص داده است و همچنین پیش بینی مطلوبی برای این مسیر وجود دارد.



شکل ۴-۷۰ نمودارهای رگرسیون خطی برای داده های آموزش



شکل ۴-۷۱ نمودارهای رگرسیون خطی برای داده های آزمایش

در شکل ۴-۷۰ و ۷۱ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. محور عمودی این نمودارها حجم تردد برداشت شده و محور افقی حجم تردد پیش‌بینی شده توسط مدل می‌باشد. همانطور که در شکل‌های مشخص می‌باشد داده‌های آزمایشی برای ۳۰۰۰ روز و داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی برخوردار هستند.

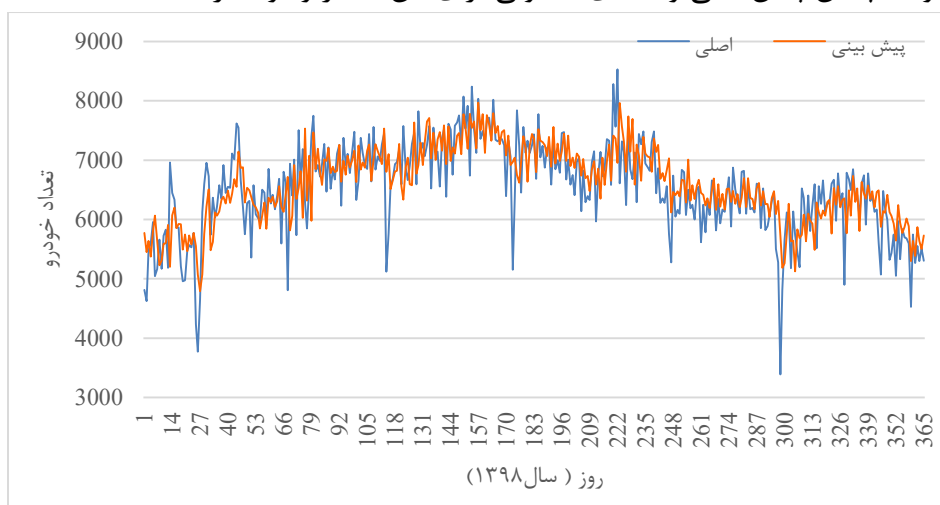
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۴۱ آمده است.

جدول ۴-۴۱ مقادیر خطا در محور مشهود - سرخس

خطا	شاخص
۰,۰۰۸	MSE
۰,۰۹۱	RMSE
۵,۶	MAPE%

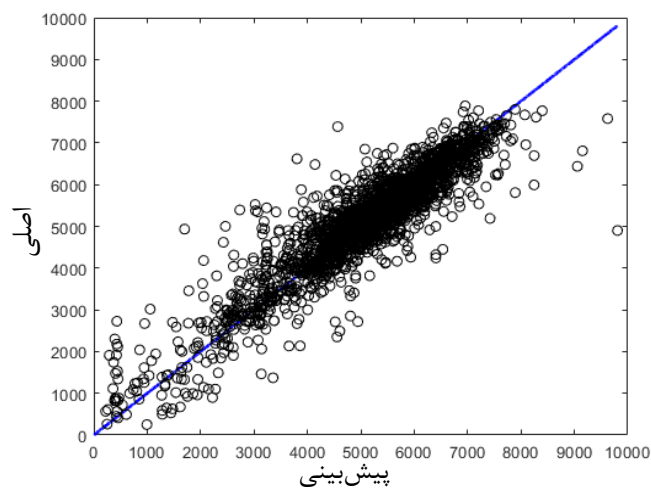
مسیر سرخس - مشهود

نمودار داده‌های حجم تردد روزانه نسبت به زمان برای ۳۳۱۶ روز در شکل ۴-۹ ترسیم شده است. در شکل ۴-۷۲ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، مدل وابستگی‌های فصلی و غیر خطی را تشخیص داده است و همچنین پیش‌بینی و خطای مطلوبی برای این مسیر وجود دارد.

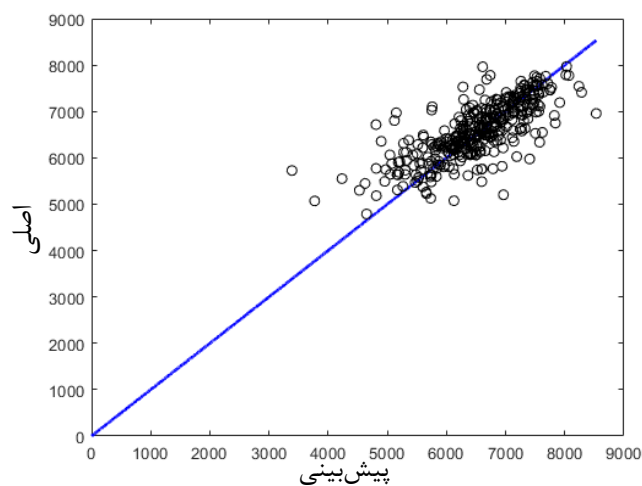


شکل ۴-۷۲ نمودار تغییرات حجم تردد نسبت به زمان در محور سرخس - مشهود

شکل ۴-۷۳ و ۷۴ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در شکل مشخص می‌باشد داده‌های آزمایشی برای ۳۰۰۰ روز و داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی برخوردار هستند.



شکل ۴-۷۳ نمودار رگرسیون خطی داده‌های آموزش



شکل ۴-۷۴ نمودار رگرسیون خطی داده‌های آزمایش

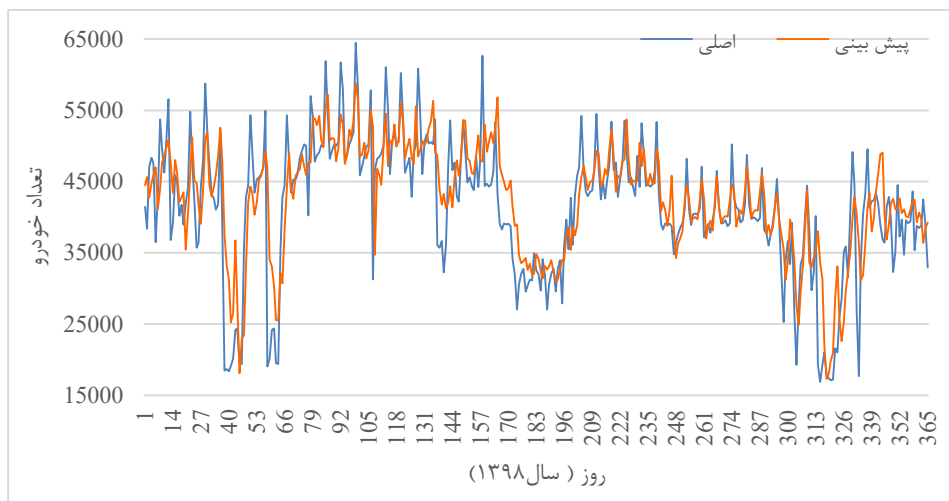
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۴۲ آمده است.

جدول ۴-۴۲ مقادیر خطا در محور سرخس - مشهد

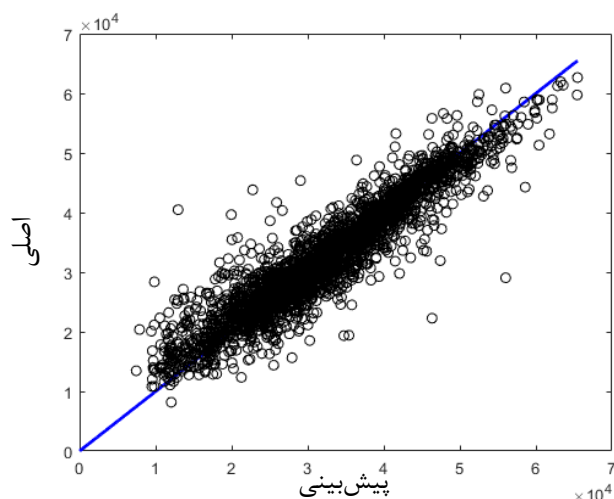
خطا	شاخص
۰,۰۱۱	MSE
۰,۱۰۷	RMSE
۶	MAPE%

مسیر مشهد - چناران

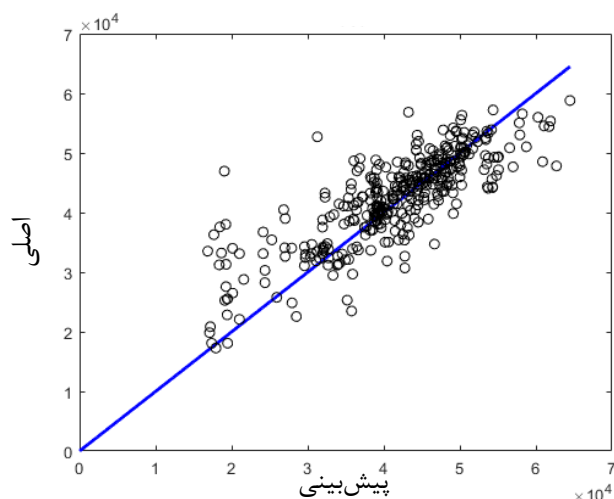
نمودار داده‌های حجم تردد روزانه نسبت به زمان برای ۳۶۵۲ روز در شکل ۴-۱۱ ترسیم شده است. در شکل ۴-۷۵ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، مدل وابستگی‌های فصلی و غیر خطی را تشخیص داده است و همچنین پیش‌بینی و خطای مطلوبی برای این مسیر وجود دارد.



شکل ۴-۷۵ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - چناران



شکل ۴-۷۶ نمودار رگرسیون خطی داده‌های آموزش



شکل ۴-۷۷ نمودار رگرسیون خطی داده‌های آزمایش

شکل ۴-۷۶ و ۷۷ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در شکل مشخص می‌باشد داده‌های آزمایشی برای ۳۳۰۰ روز و داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی برخوردار هستند.

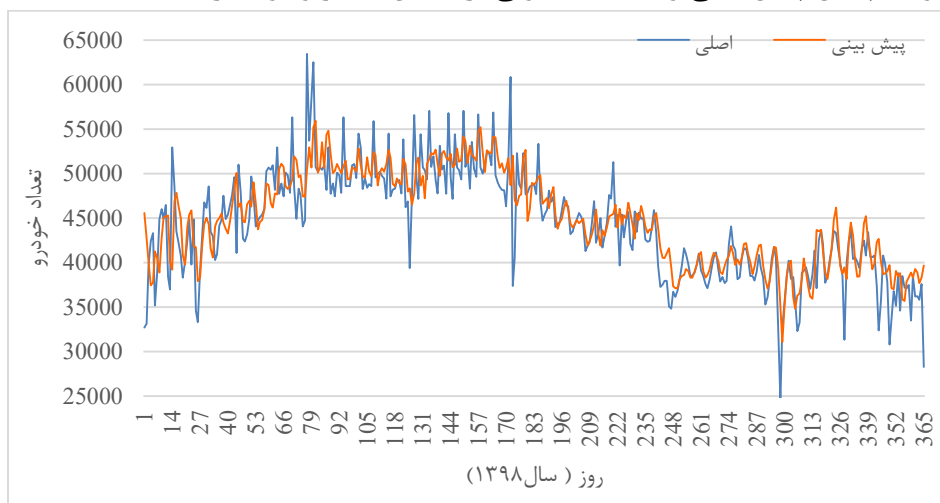
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۴۳ آمده است.

جدول ۴-۴۳ مقادیر خطا در محور مشهد - چناران

خطا	شاخص
۰,۰۳۷	MSE
۰,۱۹۳	RMSE
۱۱,۶	MAPE%

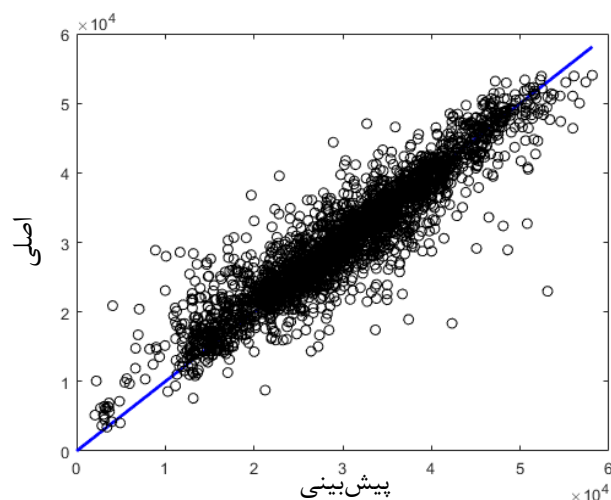
مسیر چناران - مشهد

نمودار داده‌های حجم تردد روزانه نسبت به زمان برای ۳۶۵۲ روز در شکل ۴-۱۳ ترسیم شده است. در شکل ۴-۷۸ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، مدل وابستگی‌های فصلی و غیر خطی را تشخیص داده است و همچنین پیش‌بینی و خطای مطلوبی برای این مسیر وجود دارد.

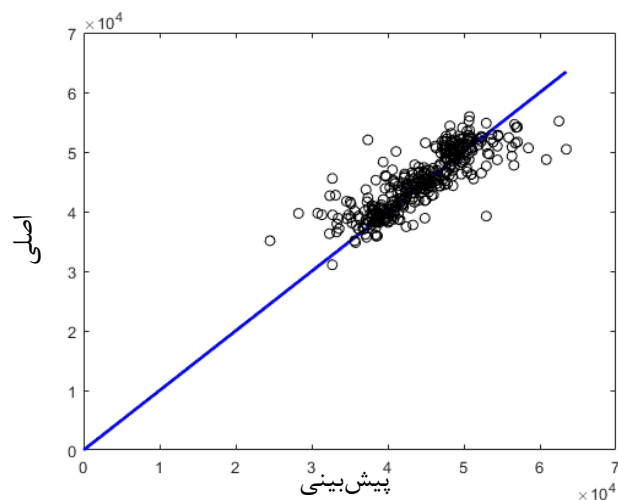


شکل ۴-۷۸ نمودار تغییرات حجم تردد نسبت به زمان در محور چناران - مشهد

شکل ۴-۷۹ و ۸۰ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در شکل مشخص می‌باشد داده‌های آزمایشی برای ۳۳۰۰ روز و داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی برخوردار هستند.



شکل ۴-۷۹ نمودار رگرسیون خطی داده‌های آموزش



شکل ۴-۸۰ نمودار رگرسیون خطی داده‌های آزمایش

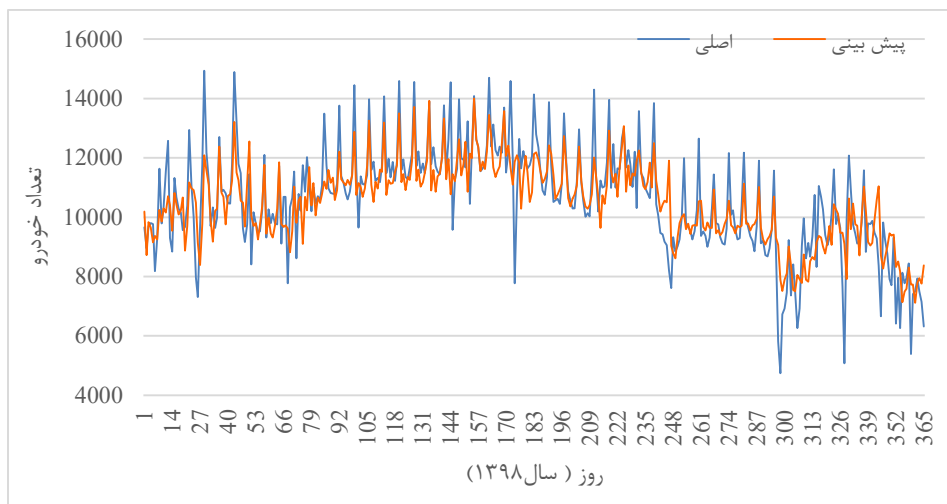
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۴۴ آمده است.

جدول ۴-۴۴ مقادیر خطا در محور چنان‌ران - مشهد

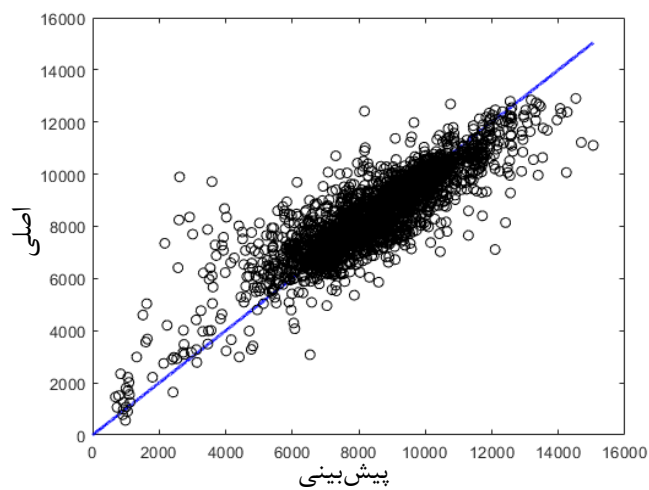
خطا	شاخص
MSE	۰,۰۱۲
RMSE	۰,۱۱
MAPE%	۵,۵

مسیر مشهد - فریمان

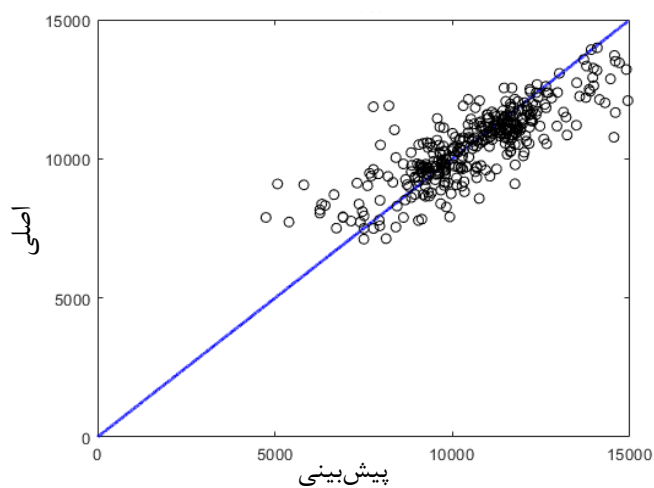
نمودار داده‌های حجم تردد روزانه نسبت به زمان برای ۳۳۱۶ روز در شکل ۴-۱۵ ترسیم شده است. در شکل ۴-۸۱ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، مدل وابستگی‌های فصلی و غیر خطی را تشخیص داده است و همچنین پیش‌بینی و خطای مطلوبی برای این مسیر وجود دارد.



شکل ۴-۸۱ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - فریمان



شکل ۴-۸۲ نمودار رگرسیون خطی داده‌های آموزش



شکل ۴-۸۳ نمودار رگرسیون خطی داده‌های آزمایش

شکل ۴-۸۲ و ۸۳ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در شکل مشخص می‌باشد داده‌های آزمایشی برای ۳۰۰۰ روز و داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی برخوردار هستند.

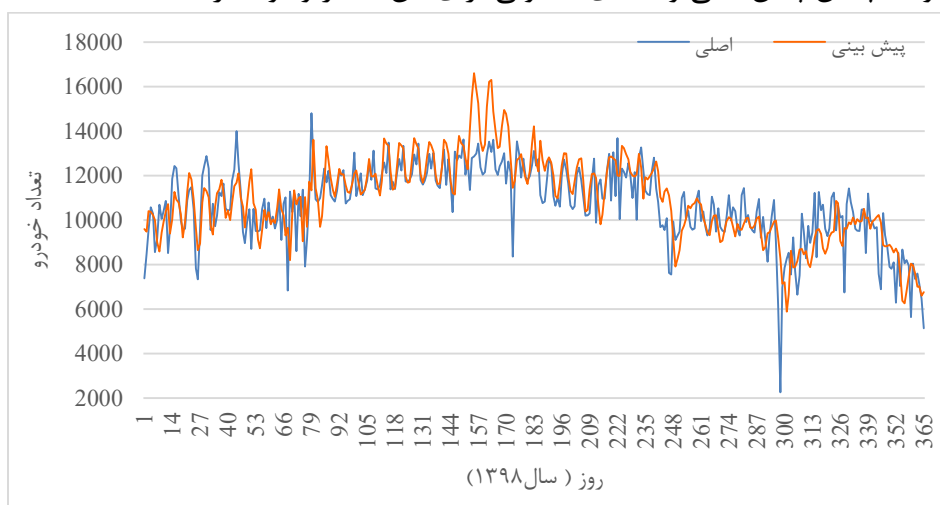
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۴۵ آمده است.

جدول ۴-۴۵ مقادیر خطا در محور مشهد - فریمان

خطا	شاخص
۰,۰۲	MSE
۰,۱۴۲	RMSE
۷,۵	MAPE%

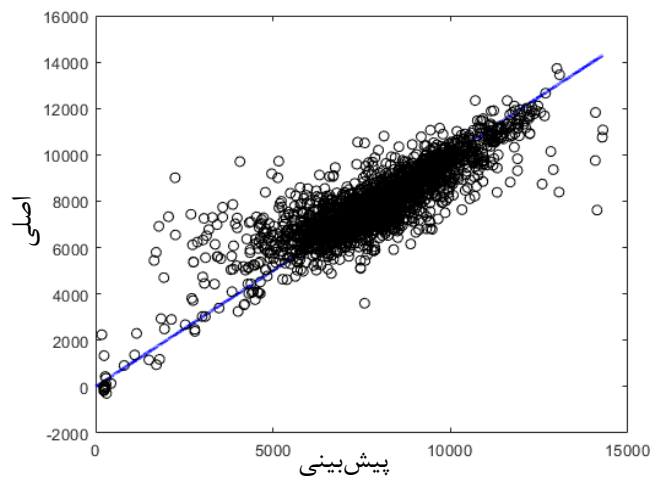
مسیر فریمان - مشهد

نمودار داده‌های حجم تردد روزانه نسبت به زمان برای ۳۳۱۶ روز در شکل ۴-۱۷ ترسیم شده است. در شکل ۴-۸۴ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، مدل وابستگی‌های فصلی و غیر خطی را تشخیص داده است و همچنین پیش‌بینی و خطای مطلوبی برای این مسیر وجود دارد.

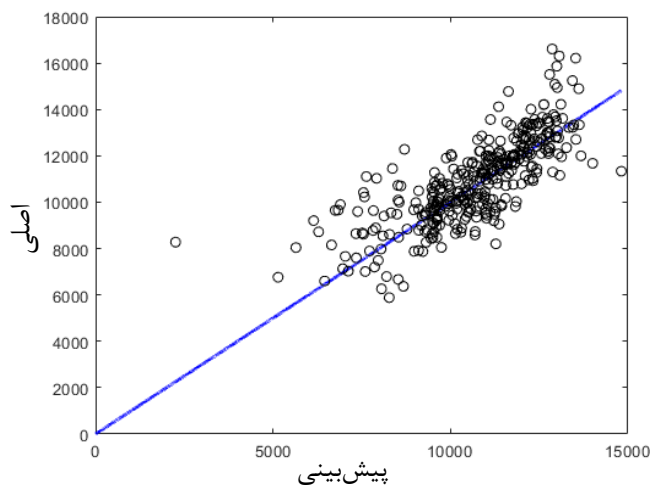


شکل ۴-۸۴ نمودار تغییرات حجم تردد نسبت به زمان در محور فریمان - مشهد

شکل ۴-۸۵ و ۸۶ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در شکل مشخص می‌باشد داده‌های آزمایشی برای ۳۰۰۰ روز و داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی برخوردار هستند.



شکل ۴-۸۵ نمودار رگرسیون خطی داده‌های آموزش



شکل ۴-۸۶ نمودار رگرسیون خطی داده‌های آزمایش

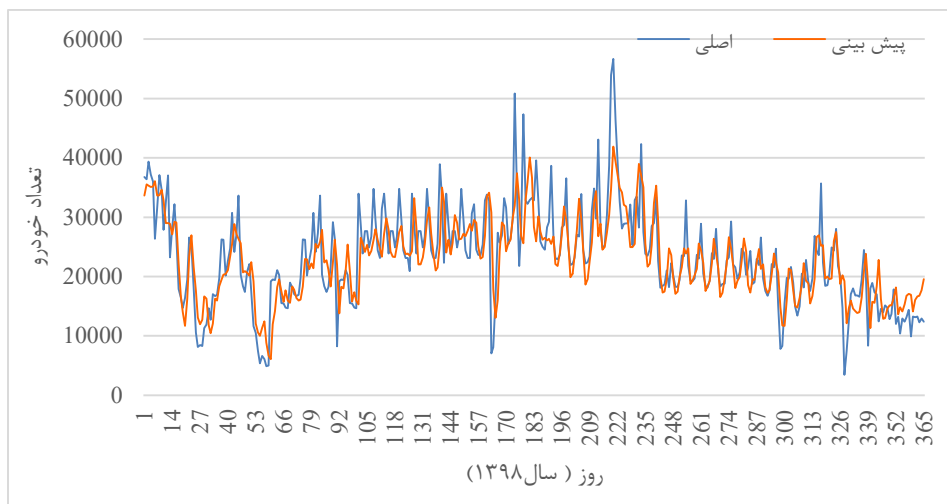
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۴۶ آمده است.

جدول ۴-۴۶ مقادیر خطا در محور فریمان - مشهد

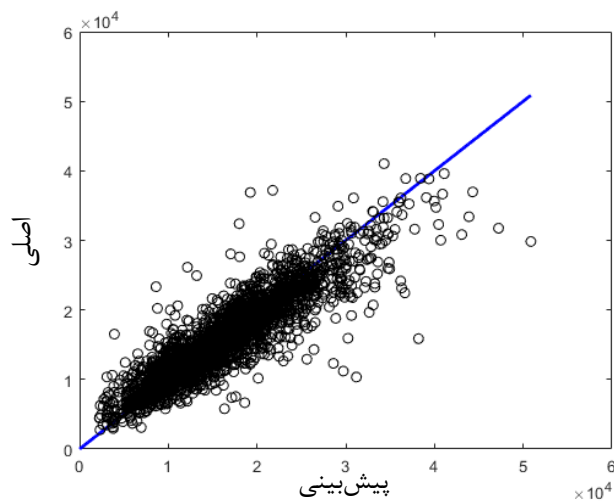
خطا	شاخص
۰,۰۲۸	MSE
۰,۱۶۸	RMSE
۹,۶	MAPE%

مسیر مشهد - ملک‌آباد

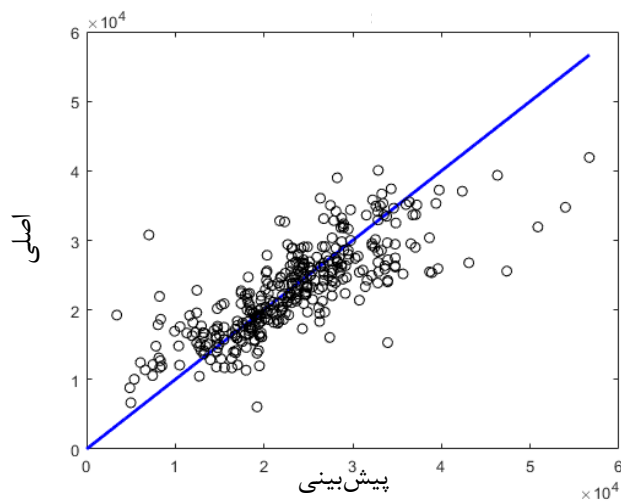
نمودار داده‌های حجم تردد روزانه نسبت به زمان برای ۳۶۵۲ روز در شکل ۴-۱۹ ترسیم شده است. در شکل ۴-۸۷ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، مدل وابستگی‌های فصلی و غیر خطی را تشخیص داده است و همچنین پیش‌بینی و خطای مطلوبی برای این مسیر وجود دارد.



شکل ۴-۸۷ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد – ملک آباد



شکل ۴-۸۸ نمودار رگرسیون خطی داده‌های آموزش



شکل ۴-۸۹ نمودار رگرسیون خطی داده‌های آزمایش

شکل ۴-۸۸ و ۸۹ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در شکل مشخص می‌باشد داده‌های آزمایشی برای ۳۳۰۰ روز و داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی برخوردار هستند.

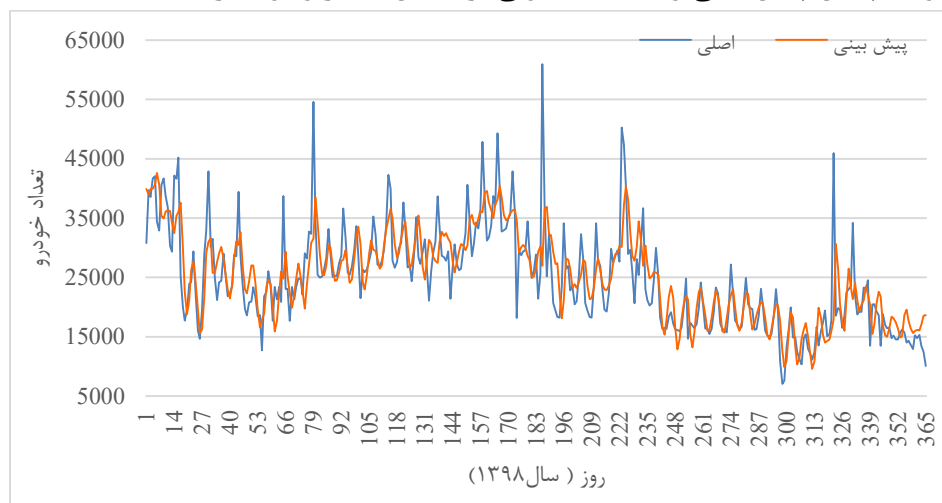
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۴۷ آمده است.

جدول ۴-۴۷ مقادیر خطا در محور مشهد - ملک آباد

خطا	شاخص
۰,۰۳۴	MSE
۰,۱۸۶	RMSE
۱۹,۱	MAPE%

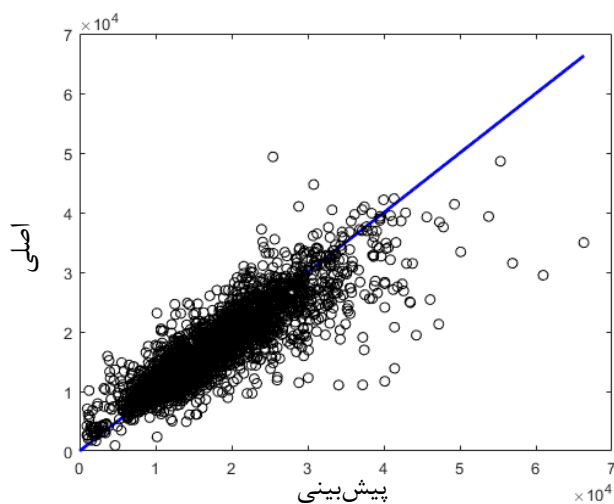
مسیر ملک آباد - مشهد

نمودار داده‌های حجم تردد روزانه نسبت به زمان برای ۳۶۵۲ روز در شکل ۴-۲۱ ترسیم شده است. در شکل ۴-۹۰ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، مدل وابستگی‌های فصلی و غیر خطی را تشخیص داده است و همچنین پیش‌بینی و خطای مطلوبی برای این مسیر وجود دارد.

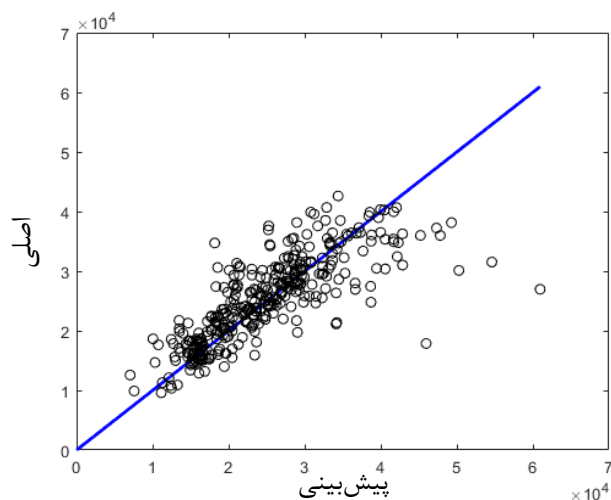


شکل ۴-۹۰ نمودار تغییرات حجم تردد نسبت به زمان در محور ملک آباد - مشهد

شکل ۴-۹۱ و ۹۲ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در شکل مشخص می‌باشد داده‌های آزمایشی برای ۳۳۰۰ روز و داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی برخوردار هستند.



شکل ۹۱-۴ نمودار رگرسیون خطی داده‌های آموزش



شکل ۹۲-۴ نمودار رگرسیون خطی داده‌های آزمایش

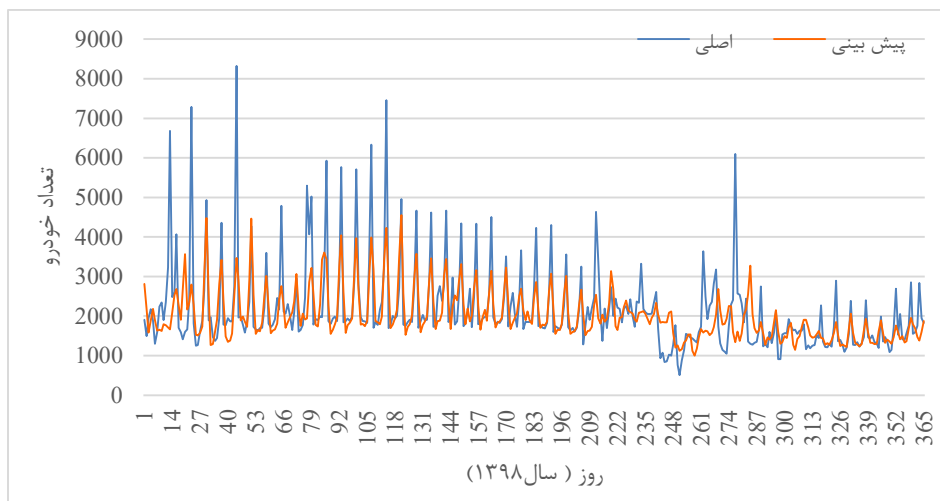
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۴۸ آمده است.

جدول ۴-۴۸ مقادیر خطا در محور ملک‌آباد - مشهد

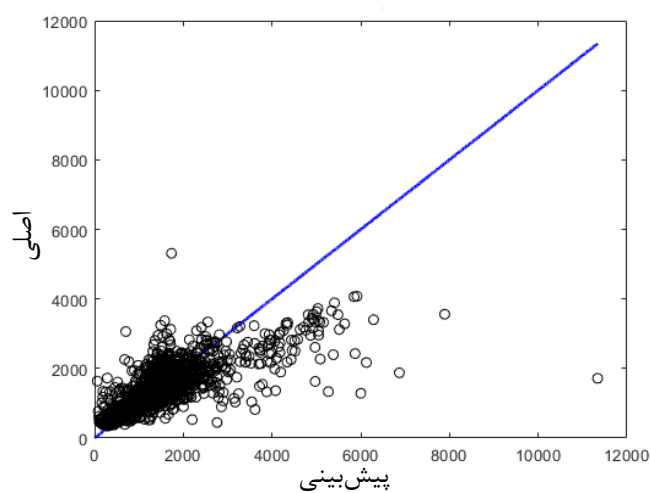
خطا	شاخص
MSE	۰,۰۲۵
RMSE	۰,۱۵۹
MAPE%	۱۴,۴

مسیر مشهد - سدکارده

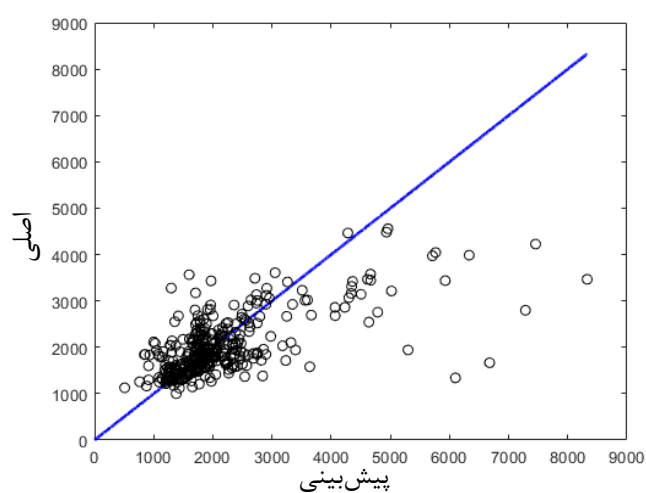
نمودار داده‌های حجم تردد روزانه نسبت به زمان برای ۳۱۳۲ روز در شکل ۴-۲۳ ترسیم شده است. در شکل ۴-۹۳ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، مدل وابستگی‌های فصلی و غیر خطی را تشخیص داده است و همچنین پیش‌بینی و خطای مطلوبی برای این مسیر وجود دارد.



شکل ۹۳-۴ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد – سدکارده



شکل ۹۴-۴ نمودار رگرسیون خطی داده‌های آموزش



شکل ۹۵-۴ نمودار رگرسیون خطی داده‌های آزمایش

شکل ۴-۹۴ و ۹۵ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در شکل مشخص می‌باشد داده‌های آزمایشی برای ۲۷۶۷ روز و داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی برخوردار هستند.

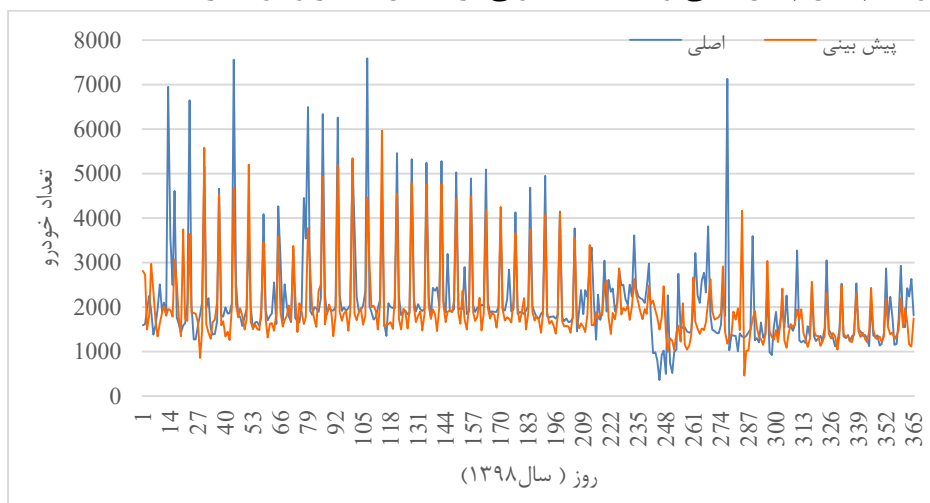
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۴۹ آمده است.

جدول ۴-۴۹ مقادیر خطا در محور مشهد - سدکارده

خطا	شاخص
۰,۰۲۱	MSE
۰,۱۴۶	RMSE
۲۱,۸	MAPE%

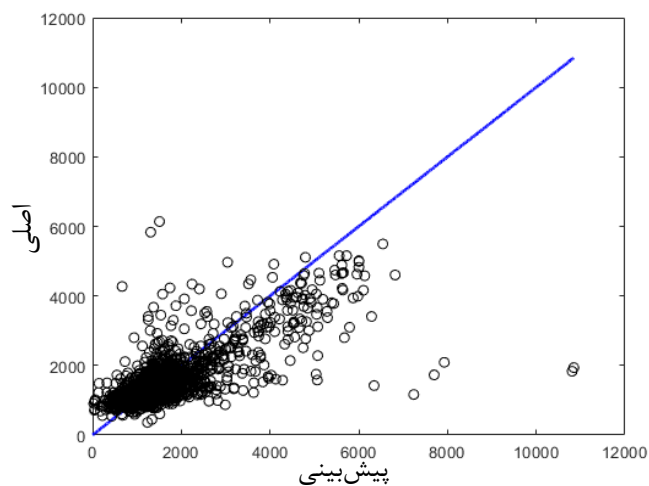
مسیر سدکارده - مشهد

نمودار داده‌های حجم تردد روزانه نسبت به زمان برای ۳۱۳۲ روز در شکل ۴-۲۵ ترسیم شده است. در شکل ۴-۹۶ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، مدل وابستگی‌های فصلی و غیر خطی را تشخیص داده است و همچنین پیش‌بینی و خطای مطلوبی برای این مسیر وجود دارد.

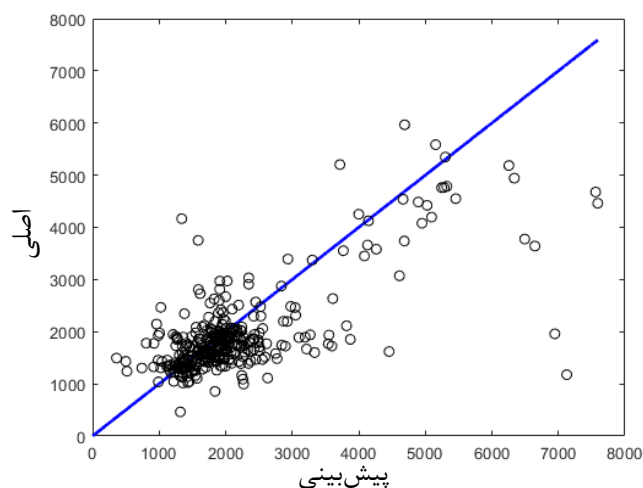


شکل ۴-۹۶ نمودار تغییرات حجم تردد نسبت به زمان در محور سدکارده - مشهد

شکل ۴-۹۷ و ۹۸ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در شکل مشخص می‌باشد داده‌های آزمایشی برای ۲۷۶۷ روز و داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی برخوردار هستند.



شکل ۹۷-۴ نمودار رگرسیون خطی داده‌های آموزش



شکل ۹۸-۴ نمودار رگرسیون خطی داده‌های آزمایش

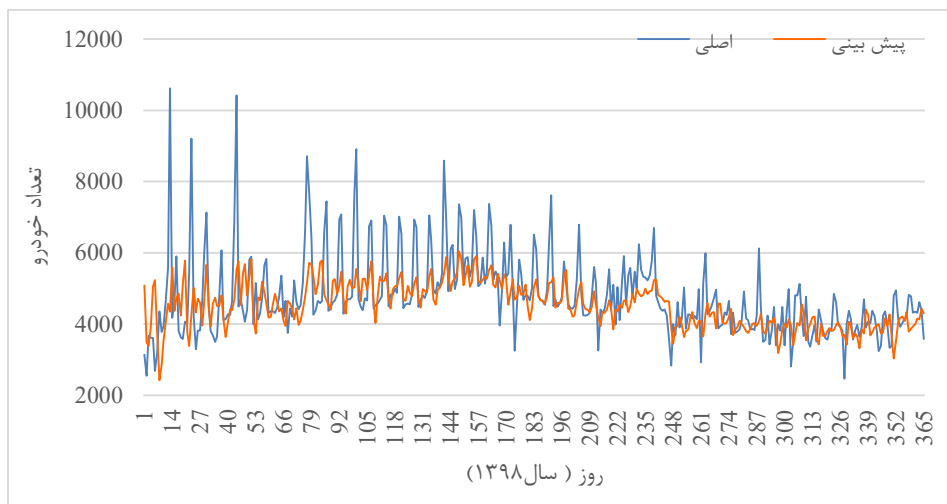
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۵۰ آمده است.

جدول ۴-۵۰ مقادیر خطا در محور سداکارده - مشهد

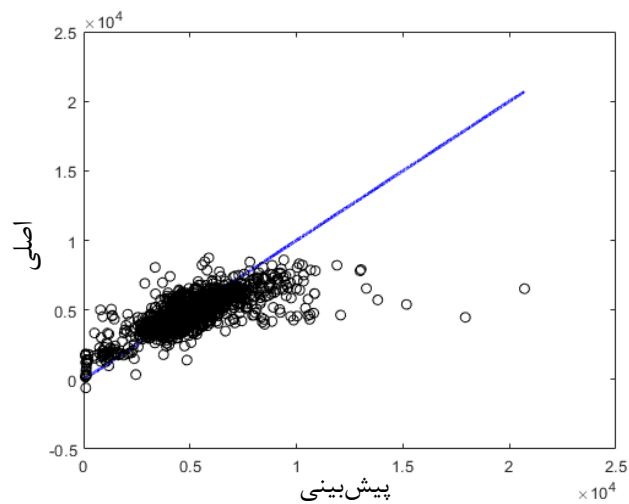
خطا	شاخص
MSE	۰,۰۲۱
RMSE	۰,۱۴۶
MAPE%	۲۲,۸

مسیر مشهد - کلات

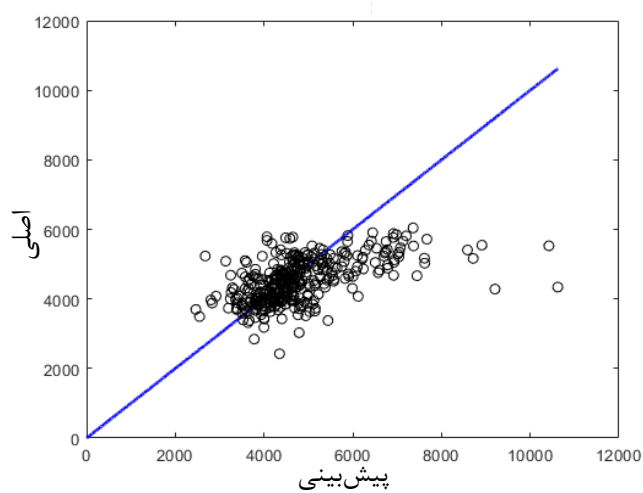
نمودار داده‌های حجم تردد روزانه نسبت به زمان برای ۲۹۵۴ روز در شکل ۴-۲۷ ترسیم شده است. در شکل ۴-۹۹ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، مدل وابستگی‌های فصلی و غیر خطی را تشخیص داده است و همچنین پیش‌بینی و خطای مطلوبی برای این مسیر وجود دارد.



شکل ۹۹-۴ نمودار تغییرات حجم تردد نسبت به زمان در محور مشهد - کلات



شکل ۱۰۰-۴ نمودار رگرسیون خطی داده‌های آموزش



شکل ۱۰۱-۴ نمودار رگرسیون خطی داده‌های آزمایش

شکل ۴-۹۷ و ۹۸ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در شکل مشخص می‌باشد داده‌های آزمایشی برای ۲۵۹۰ روز و داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی برخوردار هستند.

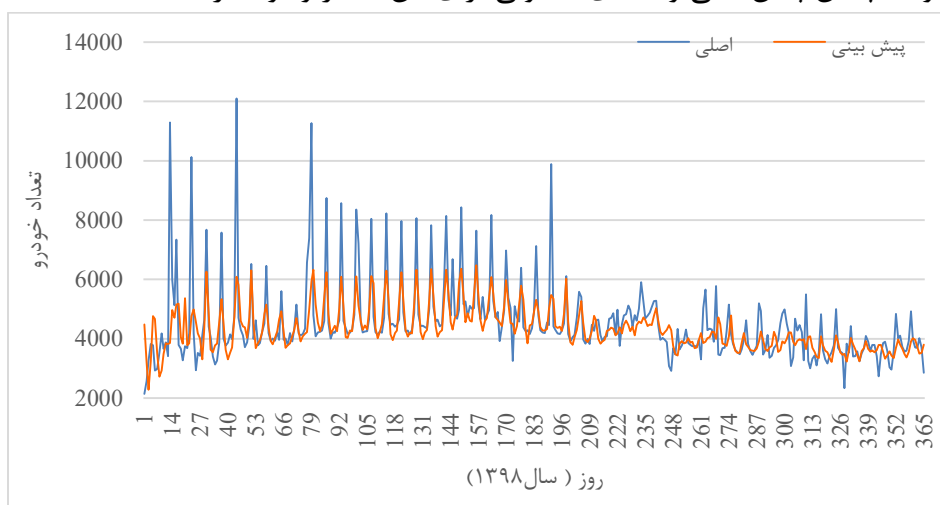
مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۵۱ آمده است.

جدول ۴-۵۱ مقادیر خطا در محور مشهد - کلات

خطا	شاخص
۰,۰۰۹	MSE
۰,۰۹۶	RMSE
۱۳,۲	MAPE%

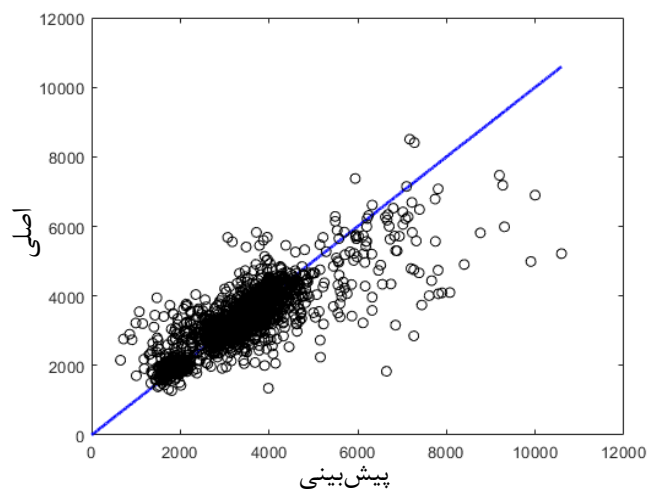
مسیر کلات - مشهد

نمودار داده‌های حجم تردد روزانه نسبت به زمان برای ۲۹۵۴ روز در شکل ۴-۲۹ ترسیم شده است. در شکل ۴-۱۰۲ نمودار تغییرات حجم تردد نسبت به زمان برای داده‌های اصلی و پیش‌بینی سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، مدل وابستگی‌های فصلی و غیر خطی را تشخیص داده است و همچنین پیش‌بینی و خطای مطلوبی برای این مسیر وجود دارد.

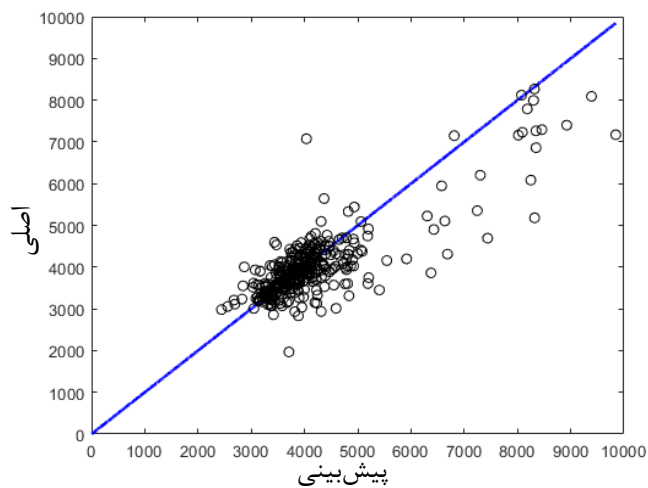


شکل ۴-۱۰۲ نمودار تغییرات حجم تردد نسبت به زمان در محور کلات - مشهد

شکل ۴-۱۰۳ و ۱۰۴ نمودارهای رگرسیون خطی برای داده آموزش و آزمایش آمده است. همانطور که در شکل مشخص می‌باشد داده‌های آزمایشی برای ۲۵۹۰ روز و داده‌های پیش‌بینی برای ۳۶۵ روز از رگرسیون خطی خوبی برخوردار هستند.



شکل ۴-۱۰۳ نمودار رگرسیون خطی داده‌های آموزش



شکل ۴-۱۰۴ نمودار رگرسیون خطی داده‌های آزمایش

. مقدار خطاهای پیش‌بینی محاسبه شده برای این محور در جدول ۴-۵۲ آمده است.

جدول ۴-۵۲ مقادیر خطا در محور کلات - مشهد

خطا	شاخص
۰,۰۱۶	MSE
۰,۱۲۸	RMSE
۱۱,۶	MAPE%

۴ - ۷ اعتبارسنجی

حال به مقایسه سه روش پیش‌بینی ارائه شده، ARIMA، شبکه عصبی مصنوعی و یادگیری عمیق LSTM پرداخته می‌شود. بدین منظور از مقایسه خطاهای پیش‌بینی MSE، RMSE و MAPE استفاده می‌شود.

۴ - ۷ - ۱ مقایسه مدل سازی ها با استفاده از خطاهای MSE و RMSE

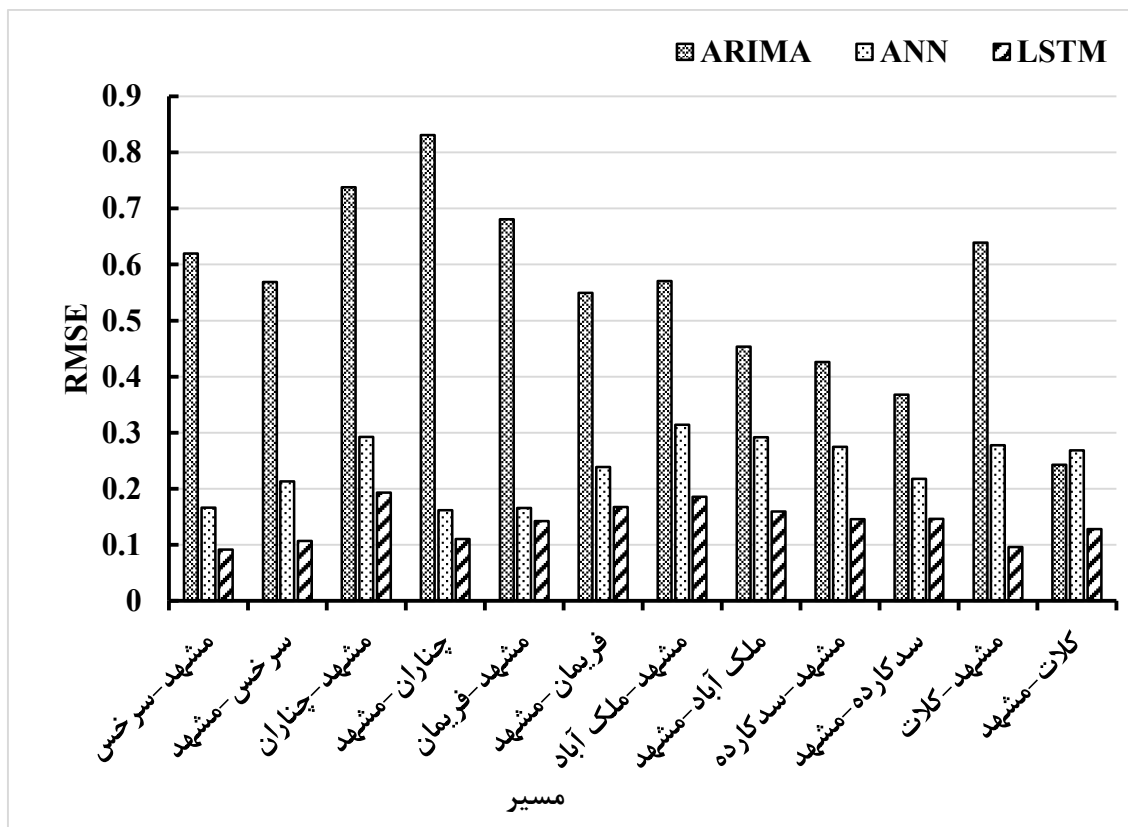
از آنجایی که سه مدل پیش بینی ARIMA، شبکه عصبی مصنوعی و یادگیری عمیق LSTM به صورت روزانه برای داده های موجود پیش بینی شده اند این سه مدل را با استفاده از خطاهای MSE و RMSE برای هر مسیر با یکدیگر مقایسه می نماییم. در جدول ۴-۵۳ برای سه روش پیش بینی برای ۱۲ مسیر موجود خطاها بررسی شده اند.

جدول ۴-۵۳ خطای MSE و RMSE برای مدلسازی روزانه سال ۱۳۹۸

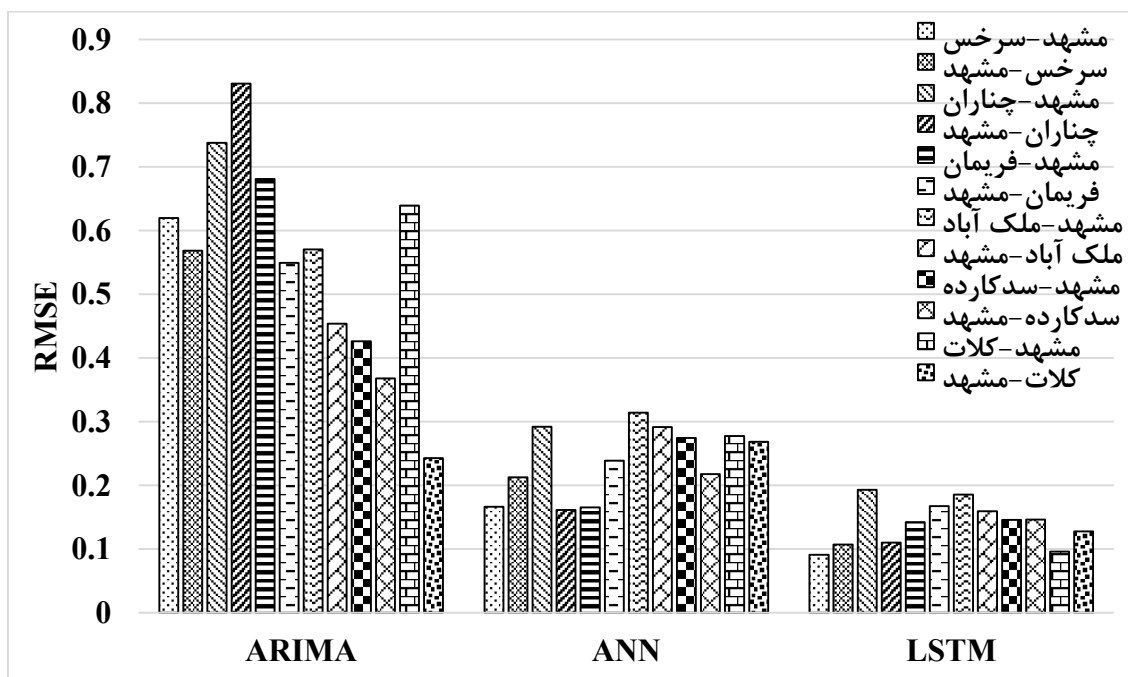
خطا		ARIMA		ANN		LSTM	
		MSE	RMSE	MSE	RMSE	MSE	RMSE
مسیر	مسیر						
مشهد-سرخس	مشهد-سرخس	۰,۳۸۴۱۸	۰,۶۱۹۸۲۲	۰,۰۲۷۶۶۹	۰,۱۶۶۳۴	۰,۰۰۸۳۳	۰,۰۹۱۲۷۱
سرخس-مشهد	سرخس-مشهد	۰,۳۲۳۲۹۲	۰,۵۶۸۵۸۸	۰,۰۴۵۳۳۹	۰,۲۱۲۹۳	۰,۰۱۱۴۶۷	۰,۱۰۷۰۸۴
مشهد-چناران	مشهد-چناران	۰,۵۴۴۰۲۵	۰,۷۳۷۵۸	۰,۰۸۵۴۵۸	۰,۲۹۲۳۳۲	۰,۰۳۷۳۰۴	۰,۱۹۳۱۴۲
چناران-مشهد	چناران-مشهد	۰,۶۹۰۲۰۶	۰,۸۳۰۷۸۷	۰,۰۲۶۰۷۴	۰,۱۶۱۴۷۴	۰,۰۱۲۱۲۷	۰,۱۱۰۱۲۳
مشهد-فریمان	مشهد-فریمان	۰,۴۶۳۵۱۳	۰,۶۸۰۸۱۸	۰,۰۲۷۴۲۹	۰,۱۶۵۶۱۷	۰,۰۲۰۲۶۳	۰,۱۴۲۳۴۸
فریمان-مشهد	فریمان-مشهد	۰,۳۰۱۶۶۱	۰,۵۴۹۲۳۷	۰,۰۵۷۱۳۷	۰,۲۳۹۰۳۳	۰,۰۲۸۰۹۳	۰,۱۶۷۶۱
مشهد-ملک آباد	مشهد-ملک آباد	۰,۳۲۵۲۸۳	۰,۵۷۰۳۳۶	۰,۰۹۸۷۴۵	۰,۳۱۴۲۳۷	۰,۰۳۴۴۲۳	۰,۱۸۵۵۳۴
ملک آباد-مشهد	ملک آباد-مشهد	۰,۲۰۵۸۸۷	۰,۴۵۳۷۴۸	۰,۰۸۵۰۸	۰,۲۹۱۶۸۵	۰,۰۲۵۳۷۲	۰,۱۵۹۲۸۶
مشهد-سدکارده	مشهد-سدکارده	۰,۱۸۱۷۲۱	۰,۴۲۶۲۸۸	۰,۰۷۵۳۷۸	۰,۲۷۴۵۵۱	۰,۰۲۱۲۸۳	۰,۱۴۵۸۸۷
سدکارده-مشهد	سدکارده-مشهد	۰,۱۳۵۴۷۶	۰,۳۶۸۰۷	۰,۰۴۷۳۴	۰,۲۱۷۵۷۸	۰,۰۲۱۴۰۳	۰,۱۴۶۲۹۸
مشهد-کلات	مشهد-کلات	۰,۴۰۸۲۷۲	۰,۶۳۸۹۶۲	۰,۰۷۶۹۹۴	۰,۲۷۷۴۷۸	۰,۰۰۹۲۵۱	۰,۰۹۶۱۸۲
کلات-مشهد	کلات-مشهد	۰,۰۵۸۹۹۲	۰,۲۴۲۸۸۳	۰,۰۷۲۲۰۹	۰,۲۶۸۷۱۷	۰,۰۱۶۳۳۹	۰,۱۲۷۸۲۴

براساس جدول ۴-۵۳ و نمودار شکل ۴-۱۰۵ به منظور مقایسه سه روش پیش بینی ARIMA، شبکه عصبی مصنوعی و یادگیری عمیق LSTM می توان دریافت که روش یادگیری عمیق LSTM برای پیش بینی حجم تردد روزانه سال ۹۸ روشی بهتر و دقیق تری می باشد و نوسانات کمتری در پیش بینی نسبت به روش شبکه عصبی مصنوعی دارد. با توجه به نمودارهای بخش ۴-۶ روش یادگیری عمیق LSTM دارای خطا کمتری می باشد و روند تغییرات و دینامیک مربوط به داده های حجم تردد روزانه، روند پیش بینی را بهتر تشخیص داده است. همچنین پیش بینی نزدیک تری به مقادیر برداشت شده حجم تردد روزانه را ارائه می دهد.

در شکل ۴-۱۰۶ نمودار خطای RMSE در سه روش پیش بینی ARIMA، شبکه عصبی مصنوعی و یادگیری عمیق LSTM مقایسه شده اند. همانطور که از شکل مشخص می باشد، روش LSTM نسبت به دو روش دیگر دیگر دارای میانگین خطای کمتری می باشد.



شکل ۴-۱۰۵ نمودار مقایسه خطای RMSE در مسیرهای موجود



شکل ۴-۱۰۶ نمودار مقایسه خطای RMSE در سه روش مدل سازی

۴ - ۷ - ۲ مقایسه مدل‌سازی‌ها با استفاده از خطا MAPE

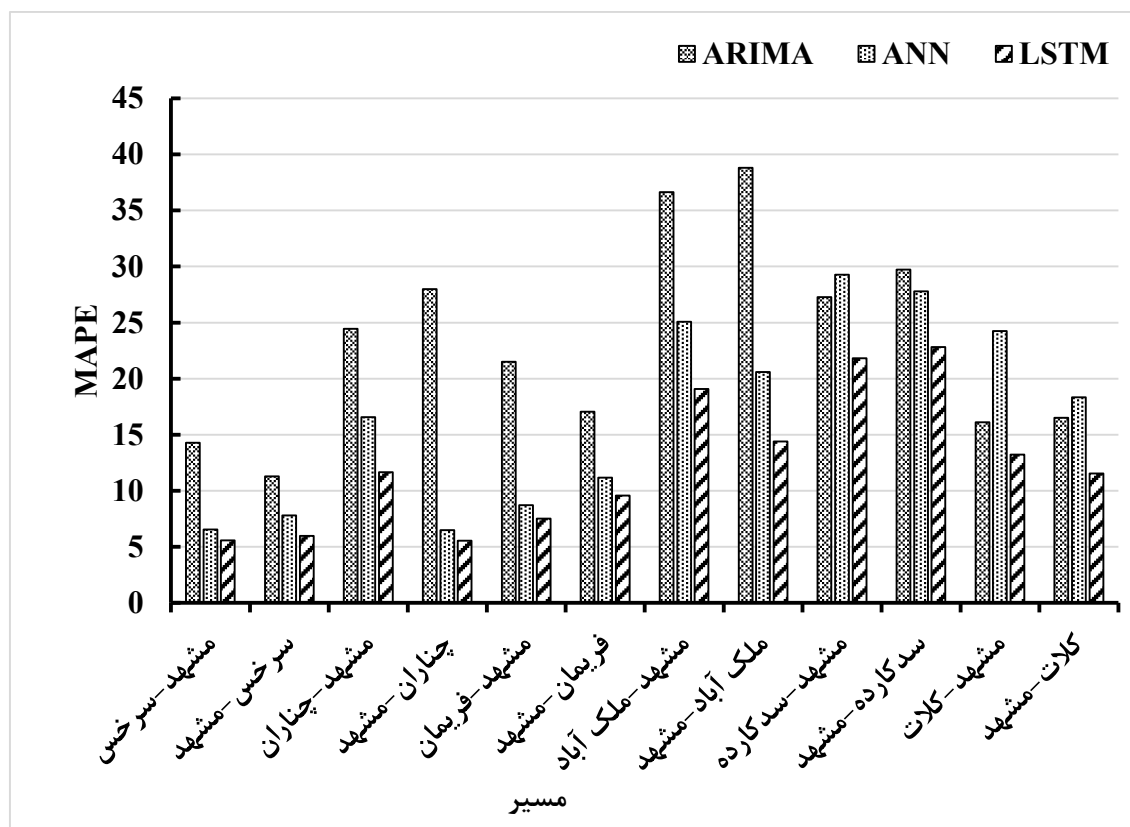
حال با استفاده از داده‌های حجم تردد روزانه سال ۹۸ برای هر مسیر در سه روش پیش‌بینی ARIMA، شبکه عصبی مصنوعی و یادگیری عمیق LSTM و میانگین درصد خطای مطلق (MAPE) می‌توان این سه مدل را باهم مقایسه نمود. در جدول ۴-۵۴ میانگین درصد خطای مطلق برای پیش‌بینی داده‌های حجم تردد ماهانه در سه روش ARIMA، شبکه عصبی مصنوعی و یادگیری عمیق LSTM بررسی شده است.

جدول ۴-۵۴ میانگین درصد خطا مطلق برای داده‌های پیش‌بینی سال ۹۸

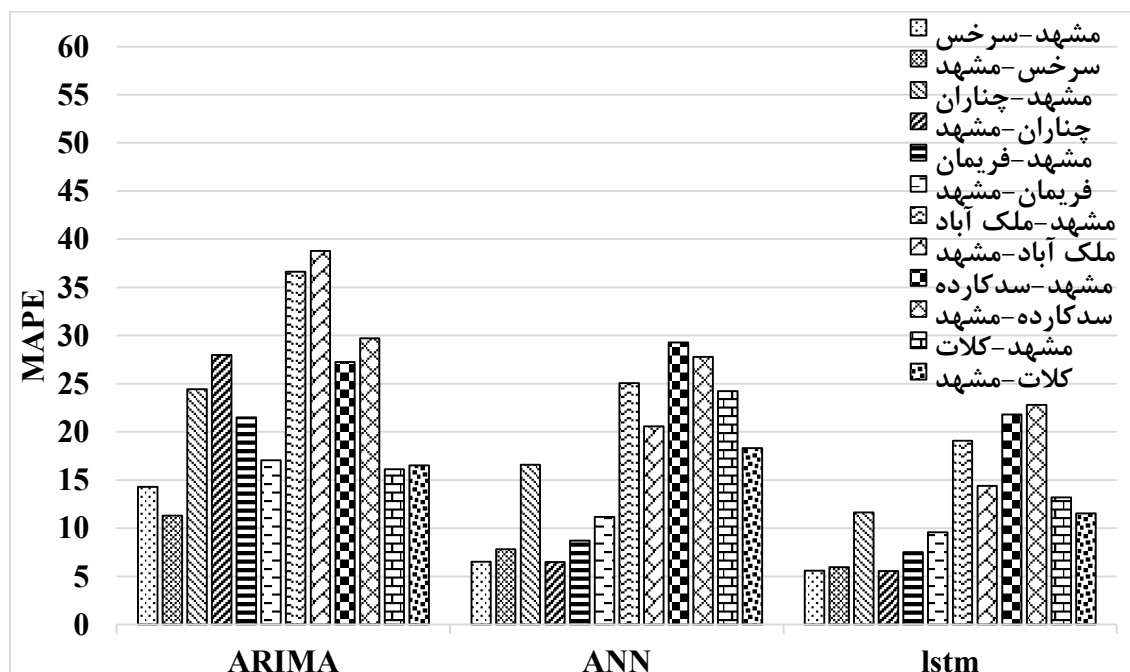
مسیر	مدل	ARIMA	ANN	LSTM
مشهد-سرخس		۱۴,۲۸۳	۶,۵۲۹۱۹۲	۵,۵۷۴۳۰۹
سرخس-مشهد		۱۱,۴۳۷۵۶	۷,۸۰۵۲۷۲	۵,۹۷۲۵۹
مشهد-چناران		۲۴,۴۳۷۵۶	۱۶,۵۷۴۵۵	۱۱,۶۴۶۲۵
چناران-مشهد		۲۷,۹۹۴۷۲	۶,۴۹۹۲۸۱	۵,۵۴۵۳۳۶
مشهد-فریمان		۲۱,۵۰۱۸۴	۸,۷۰۷۶۸۸	۷,۵۲۵۲۶
فریمان-مشهد		۱۷,۰۴۹۳۴	۱۱,۱۶۲۰۴	۹,۵۶۷۹۱۲
مشهد-ملک آباد		۳۶,۶۲۴۷۳	۲۵,۰۶۳۴۴	۱۹,۰۶۹۹۸
ملک آباد-مشهد		۳۸,۷۹۴۸۵	۲۰,۵۷۸۰۷	۱۴,۳۸۱۳۹
مشهد-سدکارده		۲۷,۲۶۰۱۵	۲۹,۲۷۶۷۶	۲۱,۸۰۳۷۳
سدکارده-مشهد		۲۹,۷۱۷۵۶	۲۷,۷۷۵۹	۲۲,۸۱۷۹۶
مشهد-کلات		۱۶,۱۱۲۱	۲۴,۲۳۵۷۲	۱۳,۲۱۲,۸۲
کلات-مشهد		۱۶,۵۱۸۰۱	۱۸,۳۳۱۵۳	۱۱,۵۵۲۸۴

براساس جدول ۴-۵۴ و نمودار شکل ۴-۱۰۷ به منظور مقایسه سه روش ARIMA، شبکه عصبی مصنوعی و یادگیری عمیق LSTM می‌توان دریافت که روش یادگیری عمیق LSTM دارای خطای مطلق کمتری نسبت به دو روش دیگر دارد.

در شکل ۴-۱۰۶ نمودار خطای MAPE در سه روش پیش‌بینی ARIMA، شبکه عصبی مصنوعی و یادگیری عمیق LSTM مقایسه شده‌اند. همانطور که از شکل مشخص می‌باشد، روش LSTM نسبت به دو روش دیگر دیگر دارای میانگین خطای کمتری می‌باشد.



شکل ۴-۱۰۷ مقایسه درصد میانگین خطای مطلق در سه روش مدل سازی



شکل ۴-۱۰۸ نمودار مقایسه خطای RMSE در سه روش مدل سازی

۴ - ۷ - ۳ مقایسه مدل سازی ها با استفاده از AADT

از آن جایی که پیش بینی های ترافیکی اغلب به صورت متوسط حجم ترافیک روزانه در سال (AADT) براساس روزهای مستند و یا مدل های پیش بینی انجام می شود، در پژوهش حاضر علاوه بر مقایسه مدل های پیش بینی و مسیرهای موجود به مقایسه AADT های این سه مدل با داده های اصلی نیز پرداخته می شود. برای این منظور داده های این بخش را در جدول ۴-۵۵ آمده است.

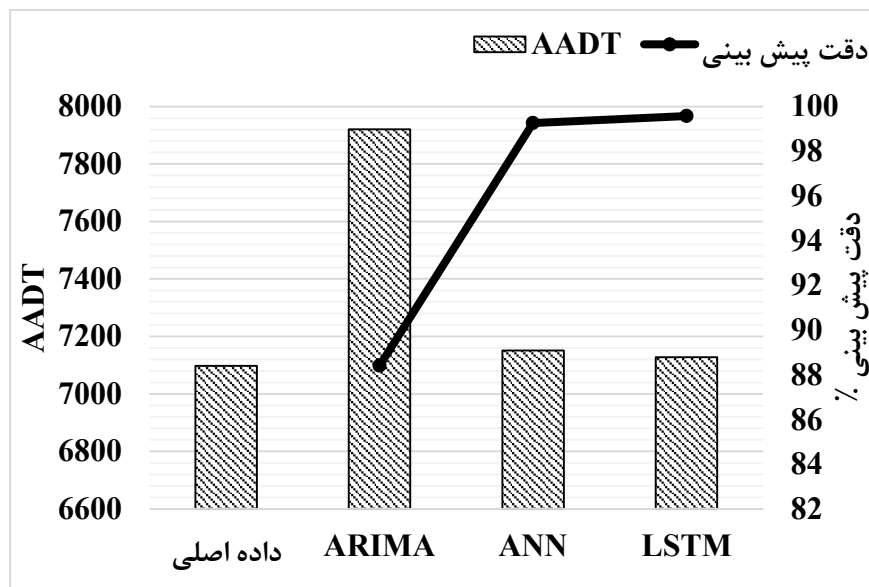
جدول ۴-۵۵ متوسط حجم ترافیک روزانه در سال ۱۳۹۸

مسیر	روش			
	LSTM	ANN	ARIMA	داده اصلی
مشهد-سرخس	۷۱۲۸	۷۲۵۱	۷۹۲۱	۷۰۹۸
سرخس-مشهد	۶۵۸۳	۶۴۰۵	۶۰۰۳	۶۵۱۷
مشهد-چناران	۴۲۴۲۱	۴۲۹۳۰	۴۵۷۲۱	۴۱۳۹۹
چناران-مشهد	۴۴۹۴۴	۴۳۸۱۴	۵۵۳۲۳	۴۴۲۷۵
مشهد-فریمان	۱۰۴۸۴	۱۰۴۹۲	۱۲۳۲۰	۱۰۵۸۰
فریمان-مشهد	۱۰۹۲۷	۱۰۴۷۲	۱۱۶۲۲	۱۰۷۰۰
مشهد-ملک آباد	۲۲۷۸۴	۲۲۵۴۵	۲۳۶۵۴	۲۳۲۱۰
ملک آباد-مشهد	۲۵۱۹۹	۲۵۶۸۲	۲۹۱۵۷	۲۴۸۴۸
مشهد-سدکارده	۲۰۱۷	۲۲۶۱	۱۹۷۰	۲۱۴۲
سدکارده-مشهد	۱۹۴۰	۲۰۵۰	۲۰۳۱	۲۱۴۶
مشهد-کلات	۴۵۲۱	۵۳۳۶	۴۷۴۵	۴۷۸۳
کلات-مشهد	۴۳۲۲	۴۲۱۸	۳۸۲۴	۴۵۰۲

حال با استفاده از داده های جدول ۴-۵۵ و درصد دقت پیش بینی ارائه شده در رابطه ۳-۶ می توان براساس AADT در مسیرهای موجود سه روش ARIMA، شبکه عصبی مصنوعی و یادگیری عمیق LSTM را باهم مقایسه نمود.

مسیر مشهد - سرخس

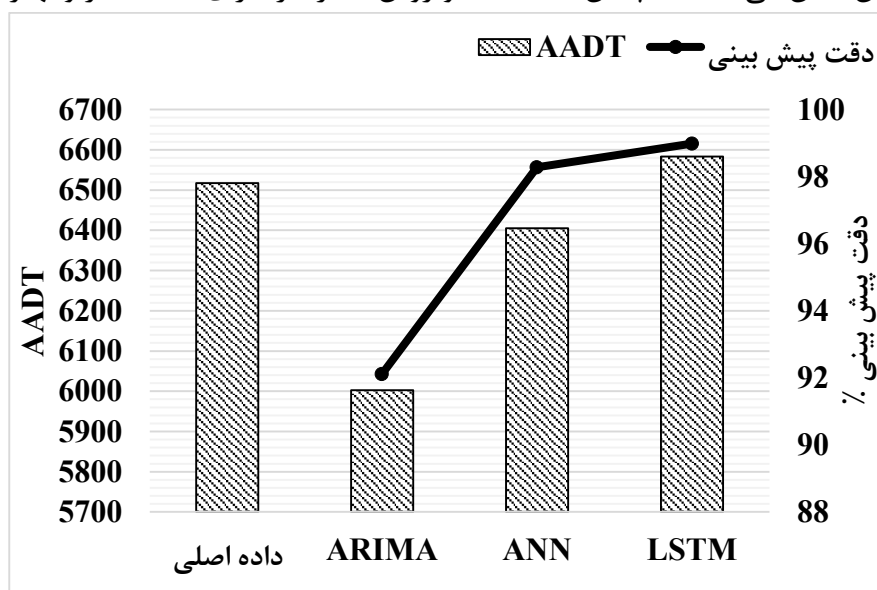
در شکل ۴-۱۰۹ نمودار ADDT برای مقایسه سه روش مدلسازی با داده های اصلی در سال ۱۳۹۸ می باشد. همانطور که در شکل مشخص می باشد، روش LSTM با ۹۹٫۵۸٪ دقت پیش بینی دقت خوبی را در این روش نشان می دهد. همچنین نسبت به دو روش دیگر نیز دارای دقت بالاتر و بهتری نیز می باشد.



شکل ۴-۱۰۹ نتایج پیش بینی AADT در محور مشهد - سرخس

مسیر سرخس - مشهد

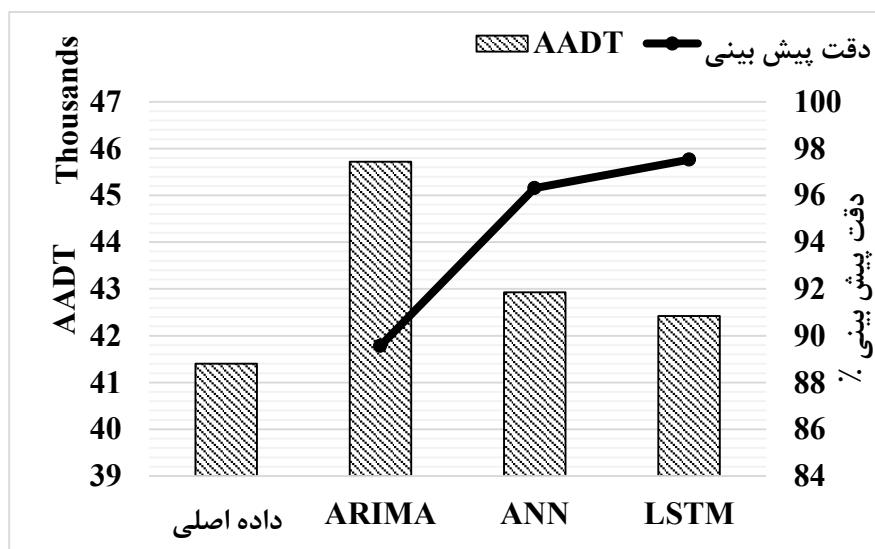
در شکل ۴-۱۱۰ نمودار ADDT برای مقایسه سه روش مدلسازی با داده‌های اصلی در سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، روش LSTM با ۹۸,۹۹٪ دقت پیش‌بینی دقت خوبی را در این روش نشان می‌دهد. همچنین نسبت به دو روش دیگر نیز دارای دقت بالاتر و بهتری می‌باشد.



شکل ۴-۱۱۰ نتایج پیش بینی AADT در محور سرخس - مشهد

مسیر مشهد - چناران

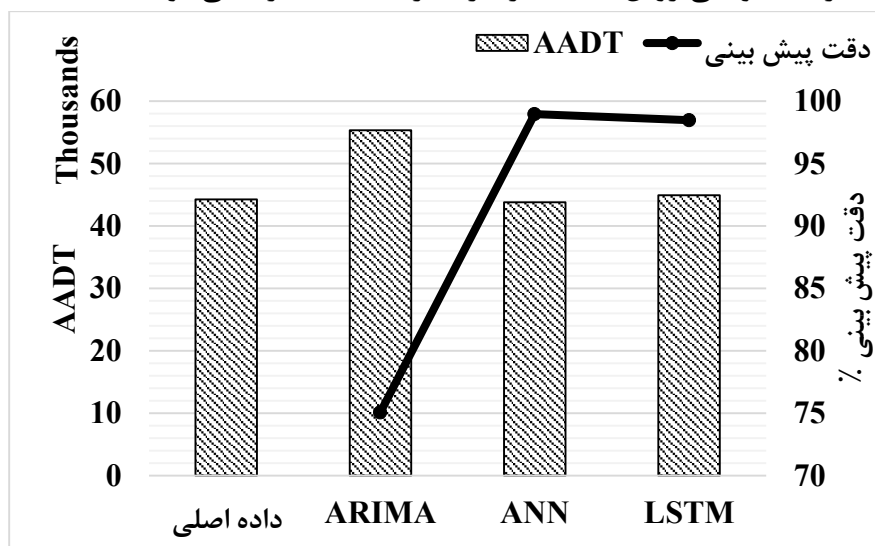
در شکل ۴-۱۱۱ نمودار ADDT برای مقایسه سه روش مدلسازی با داده‌های اصلی در سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، روش LSTM با ۹۷,۵۳٪ دقت پیش‌بینی دقت خوبی را در این روش نشان می‌دهد. همچنین نسبت به دو روش دیگر نیز دارای دقت بالاتر و بهتری می‌باشد.



شکل ۱۱۱-۴ نتایج پیش بینی AADT در محور مشهد - چناران

مسیر چناران - مشهد

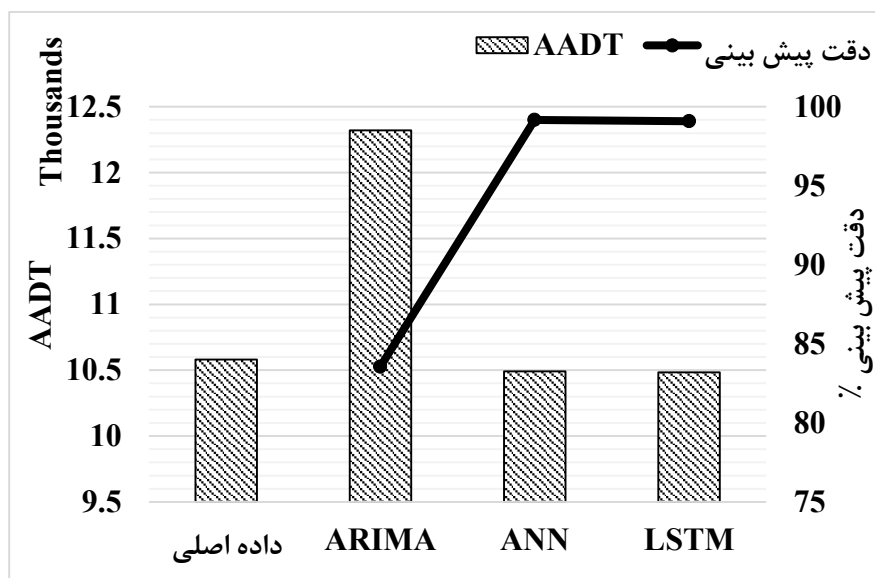
در شکل ۱۱۲-۴ نمودار ADDT برای مقایسه سه روش مدلسازی با داده‌های اصلی در سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، روش ANN با ۹۸,۹۶٪ دقت پیش‌بینی دقت خوبی را در این روش نشان می‌دهد. همچنین با روش LSTM با ۹۸,۴۹٪ نسبت به روش ANN کمتر از ۰,۵٪ اختلاف دارد که توانایی روش LSTM را در کنار ANN منجر به می‌شود.



شکل ۱۱۲-۴ نتایج پیش بینی AADT در محور سرخس - مشهد

مسیر مشهد - فریمان

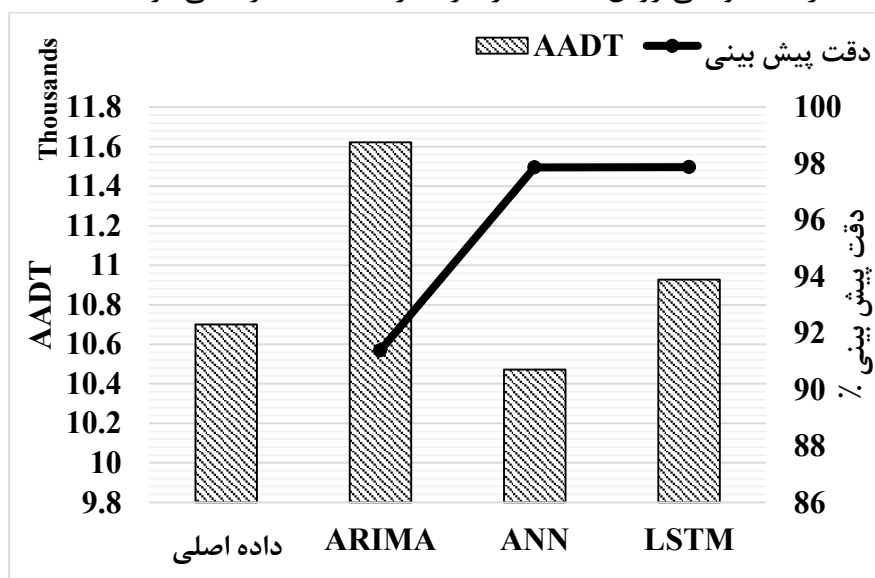
در شکل ۱۱۳-۴ نمودار ADDT برای مقایسه سه روش مدلسازی با داده‌های اصلی در سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، روش ANN با ۹۹,۱۷٪ دقت پیش‌بینی دقت خوبی را در این روش نشان می‌دهد. همچنین با روش LSTM با ۹۹,۱۰٪ نسبت به روش ANN کمتر از ۰,۱٪ اختلاف دارد که توانایی روش LSTM را در کنار ANN منجر به می‌شود.



شکل ۴-۱۱۳ نتایج پیش بینی AADT در محور مشهد - فریمان

مسیر فریمان - مشهد

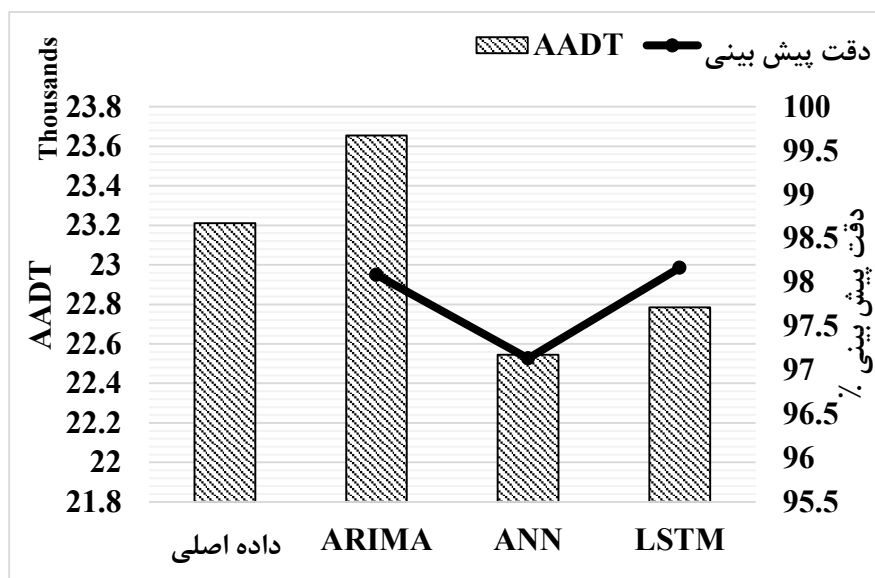
در شکل ۴-۱۱۴ نمودار ADDT برای مقایسه سه روش مدلسازی با داده‌های اصلی در سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، روش LSTM با ۹۷,۸۸٪ دقت پیش‌بینی دقت خوبی را در این روش نشان می‌دهد. همچنین با روش ANN با ۹۷,۸۷٪ نسبت به روش LSTM کمتر از ۰,۰۱٪ اختلاف دارد که توانایی روش LSTM را در کنار ANN منجر به می‌شود.



شکل ۴-۱۱۴ نتایج پیش بینی AADT در محور فریمان - مشهد

مسیر مشهد - ملک‌آباد

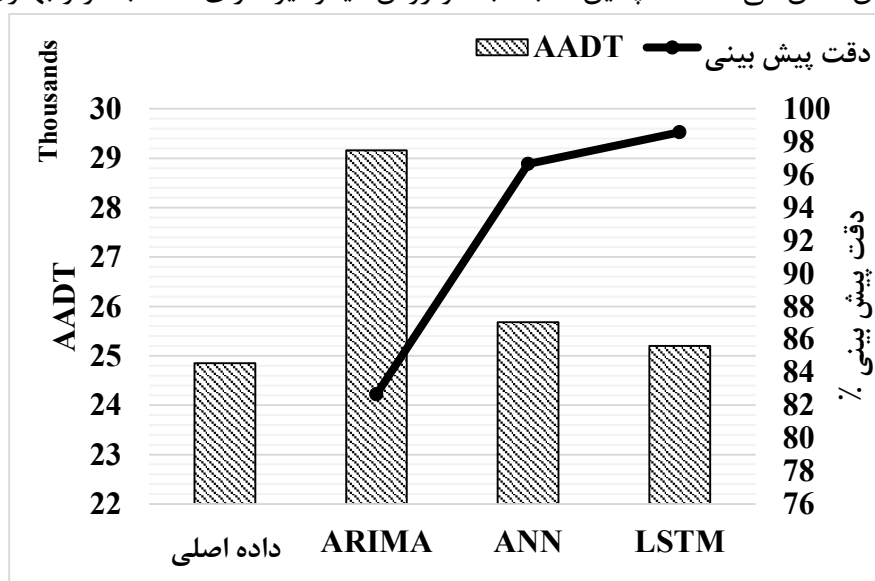
در شکل ۴-۱۱۵ نمودار ADDT برای مقایسه سه روش مدلسازی با داده‌های اصلی در سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، روش LSTM با ۹۸,۱۶٪ دقت پیش‌بینی دقت خوبی را در این روش نشان می‌دهد. همچنین نسبت به دو روش دیگر نیز دارای دقت بالاتر و بهتری می‌باشد.



شکل ۴-۱۱۵ نتایج پیش بینی AADT در محور مشهد - ملک آباد

مسیر ملک آباد - مشهد

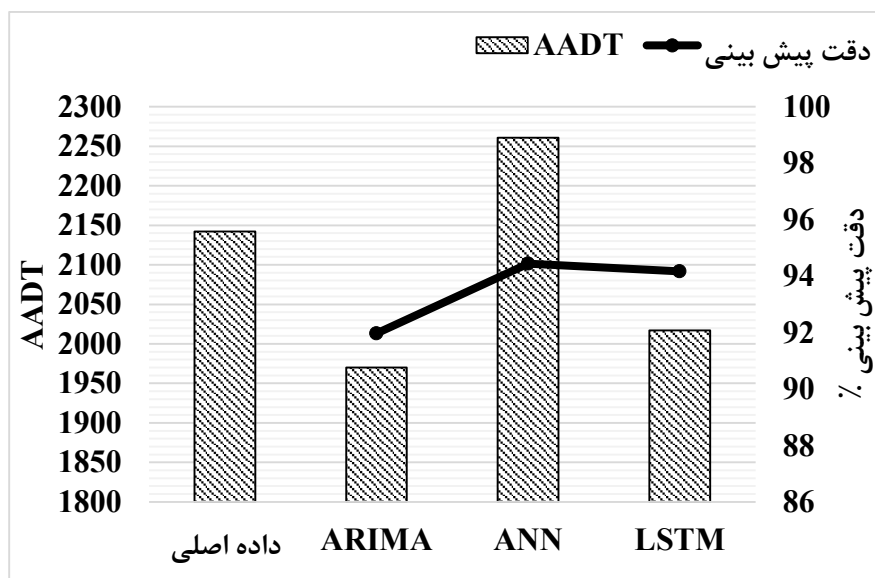
در شکل ۴-۱۱۶ نمودار ADDT برای مقایسه سه روش مدلسازی با داده‌های اصلی در سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، روش LSTM با ۹۸,۵۹٪ دقت پیش‌بینی دقت خوبی را در این روش نشان می‌دهد. همچنین نسبت به دو روش دیگر نیز دارای دقت بالاتر و بهتری می‌باشد.



شکل ۴-۱۱۶ نتایج پیش بینی AADT در محور ملک آباد - مشهد

مسیر مشهد - سدکارده

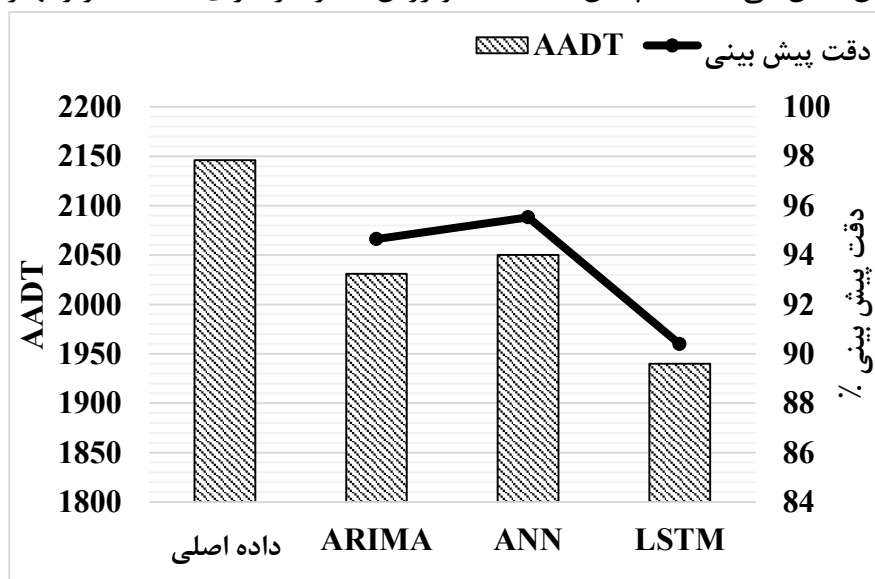
در شکل ۴-۱۱۷ نمودار ADDT برای مقایسه سه روش مدلسازی با داده‌های اصلی در سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، روش ANN با ۹۴,۴۴٪ دقت پیش‌بینی دقت خوبی را در این روش نشان می‌دهد. همچنین با روش LSTM با ۹۴,۱۶٪ نسبت به روش ANN کمتر از ۰,۳٪ اختلاف دارد که توانایی روش LSTM را در کنار ANN منجر به می‌شود.



شکل ۴-۱۱۷ نتایج پیش بینی AADT در محور مشهد - سدکارده

مسیر سدکارده - مشهد

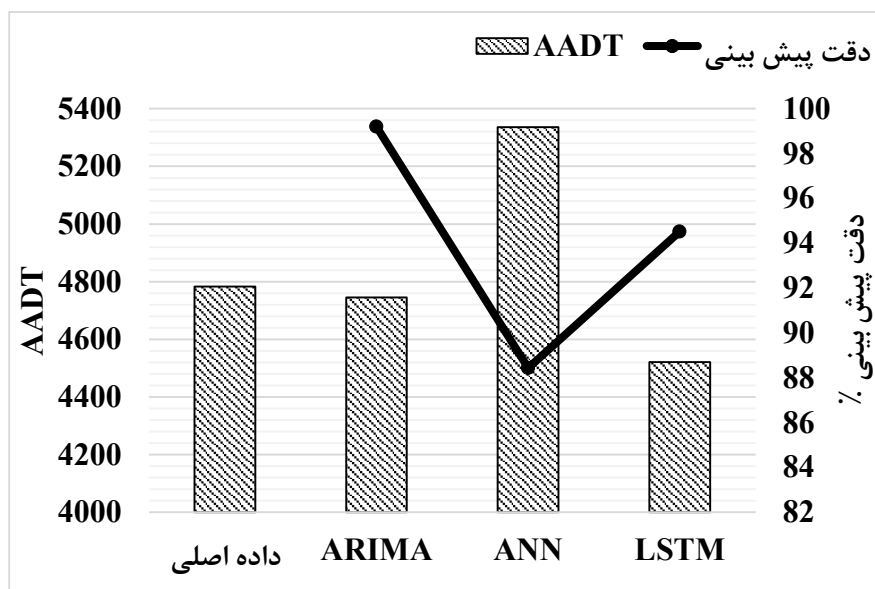
در شکل ۴-۱۱۸ نمودار ADDT برای مقایسه سه روش مدلسازی با داده‌های اصلی در سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، روش ANN با ۹۵٫۵۳٪ دقت پیش‌بینی دقت خوبی را در این روش نشان می‌دهد. همچنین نسبت به دو روش دیگر نیز دارای دقت بالاتر و بهتری می‌باشد.



شکل ۴-۱۱۸ نتایج پیش بینی AADT در محور ملک‌آباد - مشهد

مسیر مشهد - کلات

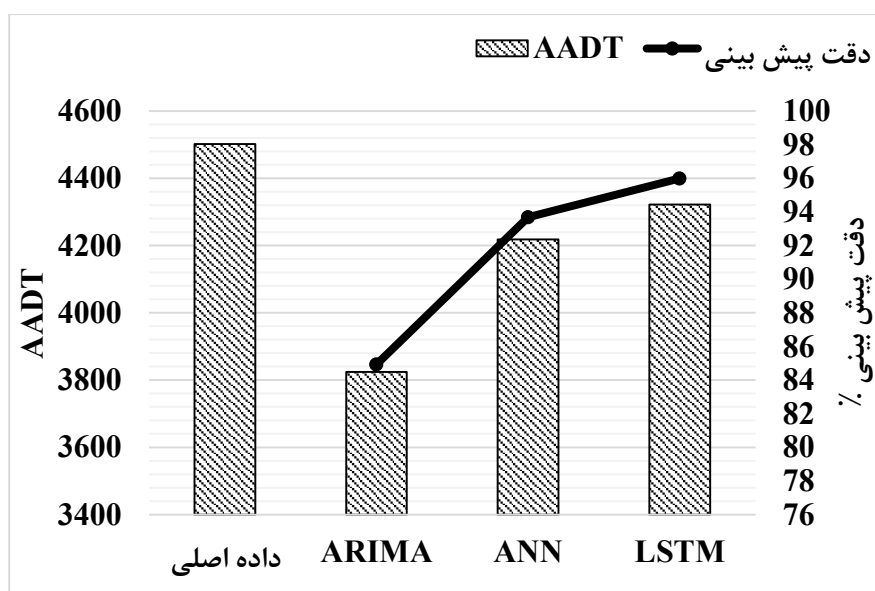
در شکل ۴-۱۱۹ نمودار ADDT برای مقایسه سه روش مدلسازی با داده‌های اصلی در سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، روش ARIMA با ۹۹٫۲۱٪ به ظاهر دقت پیش‌بینی دقت خوبی را در این روش نشان می‌دهد. اما در شکل ۴-۲۸ مشخص است که این روش اصلاً دقت مناسبی را برای پیش‌بینی نشان نمی‌دهد.



شکل ۴-۱۱۹ نتایج پیش بینی AADT در محور مشهد - کلات

مسیر کلات - مشهد

در شکل ۴-۱۲۰ نمودار ADDT برای مقایسه سه روش مدلسازی با داده‌های اصلی در سال ۱۳۹۸ می‌باشد. همانطور که در شکل مشخص می‌باشد، روش LSTM با ۹۶٪ دقت پیش‌بینی دقت خوبی را در این روش نشان می‌دهد. همچنین نسبت به دو روش دیگر نیز دارای دقت بالاتر و بهتری می‌باشد.



شکل ۴-۱۲۰ نتایج پیش بینی AADT در محور کلات - مشهد

حال با استفاده از روش‌های RMSE، MAPE و AADT ارائه شده در اعتبارسنجی برای مدل‌های پیش‌بینی ARIMA، شبکه عصبی مصنوعی و یادگیری عمیق LSTM می‌توان دریافت که روش یادگیری عمیق LSTM برای پیش‌بینی حجم تردد روزانه سال ۹۸ نسبت به دو روش دیگر از دقت و

پیش‌بینی بهتری برخوردار است در جدول ۴-۵۶ برای ۱۲ مسیر مورد ارزیابی در این پژوهش بهترین مدل پیش‌بینی برای هر کدام از روش‌های اعتبارسنجی خلاصه شده‌است.

جدول ۴-۵۶ بهترین مدل پیش‌بینی در سه روش اعتبارسنجی RMSE، MAPE و AADT

مسیر	روش		
	AADT	MAPE	RMSE
مشهد-سرخس	LSTM	LSTM	LSTM
سرخس-مشهد	LSTM	LSTM	LSTM
مشهد-چناران	LSTM	LSTM	LSTM
چناران-مشهد	ANN	LSTM	LSTM
مشهد-فریمان	ANN	LSTM	LSTM
فریمان-مشهد	LSTM	LSTM	LSTM
مشهد-ملک آباد	LSTM	LSTM	LSTM
ملک آباد-مشهد	LSTM	LSTM	LSTM
مشهد-سدکارده	ANN	LSTM	LSTM
سدکارده-مشهد	ANN	LSTM	LSTM
مشهد-کلات	ARIMA	LSTM	LSTM
کلات-مشهد	LSTM	LSTM	LSTM

در اکثر مسیرهای فوق روش LSTM در اعتبارسنجی RMSE، MAPE و AADT از صحت بالایی برخوردار بود. اما نکته قابل ذکر این است که در مسیرهای منتهی به سدکارده و کلات که جزء مسیرهای توریستی - تفریحی می باشند دقت پیش‌بینی کمتر شده است و دلیل این امر وجود داده‌های روزهای آخر هفته و روزهای تعطیل می‌باشد.

فصل ۵ : نتیجه گیری و پیشنهادات

۵ - ۱ مقدمه

روش‌های پیش‌بینی جریان ترافیک را می‌توان به سه دسته طبقه‌بندی نمود: روش‌های پارامتریک، غیرپارامتریک و ترکیبی. روش‌های ترکیبی عمدتاً ترکیب روش‌های پارامتریک و غیرپارامتریک می‌باشند. فناوری حسگرهای ترافیکی نوین، حجم زیادی از داده‌های ترافیکی را با وضوح مکانی و زمانی بالا در دسترس قرار داده‌اند. روبه‌رو شدن با این داده‌های بزرگ ترافیکی با روش‌های پارامتریک یک چالش بزرگ می‌باشد. برای روش‌های غیرپارامتری که اکثر آن‌ها از نظر معماری، سطحی می‌باشند نمی‌توانند در اطلاعات عمیق ترافیکی نفوذ کنند. در چند سال اخیر یادگیری عمیق که یک روش یادگیری ماشین در حال پیشرفت می‌باشد توجه زیادی را از طرف مهندسان ترافیک و دانشگاه به خود جلب کرده است. پیش‌بینی جریان ترافیک براساس روش یادگیری عمیق به یک روند جدید در این مبحث تبدیل شده است. مطالب این بخش

۵ - ۲ نتایج پیش‌بینی

در این بخش نتایج حاصل از پیش‌بینی در سه مدل پیش‌بینی ARIMA، شبکه عصبی مصنوعی و یادگیری عمیق LSTM بررسی می‌شود.

مدل ARIMA

در اعتبارسنجی این مدل به وسیله شاخص RMSE، کمترین خطای مربوط به محور کلات - مشهد با خطای ۰,۲۴ و بیشترین خطا مربوط محور چناران - مشهد با خطای ۰,۸۳ می‌باشد.

در اعتبارسنجی این مدل به وسیله شاخص MAPE، کمترین خطای مربوط به محور سرخس - مشهد با خطای ۱۱,۳٪ و بیشترین خطا مربوط محور ملک‌آباد - مشهد با خطای ۳۶,۸٪ می‌باشد.

در اعتبارسنجی این مدل به وسیله شاخص AADT، بهترین دقت پیش‌بینی مربوط به محور کلات - مشهد با دقت پیش‌بینی ۹۹,۲٪ و بدترین دقت پیش‌بینی مربوط محور چناران - مشهد با دقت پیش‌بینی ۷۵٪ می‌باشد.

مدل شبکه عصبی مصنوعی

در اعتبارسنجی این مدل به وسیله شاخص RMSE، کمترین خطای مربوط به محور چناران - مشهد با خطای ۰,۱۶ و بیشترین خطا مربوط محور مشهد - ملک‌آباد با خطای ۰,۳۱ می‌باشد.

در اعتبارسنجی این مدل به وسیله شاخص MAPE، کمترین خطای مربوط به محور چناران - مشهد با خطای ۶,۵٪ و بیشترین خطا مربوط محور مشهد - سدکارده با خطای ۲۹,۳٪ می‌باشد.

در اعتبارسنجی این مدل به وسیله شاخص AADT، بهترین دقت پیش‌بینی مربوط به محور مشهد - سرخس با دقت پیش‌بینی ۹۹,۳٪ و بدترین دقت پیش‌بینی مربوط به محور مشهد - کلات با دقت پیش‌بینی ۸۸,۴٪ می‌باشد.

مدل یادگیری عمیق LSTM

در اعتبارسنجی این مدل به وسیله شاخص RMSE، کمترین خطای مربوط به محور مشهد - سرخس با خطای ۰,۰۹ و بیشترین خطا مربوط به محور مشهد - چناران با خطای ۰,۱۹ می‌باشد.

در اعتبارسنجی این مدل به وسیله شاخص MAPE، کمترین خطای مربوط به محور مشهد - سرخس با خطای ۵,۵٪ و بیشترین خطا مربوط به محور سدکارده - مشهد با خطای ۲۲,۸٪ می‌باشد.

در اعتبارسنجی این مدل به وسیله شاخص AADT، بهترین دقت پیش‌بینی مربوط به محور مشهد - سرخس با دقت پیش‌بینی ۹۹,۵۸٪ و بدترین دقت پیش‌بینی مربوط به محور مشهد - کلات با دقت پیش‌بینی ۹۰,۴٪ می‌باشد.

۵ - ۳ مقایسه سه مدل پیش‌بینی

در این بخش سه مدل پیش‌بینی ARIMA، شبکه عصبی مصنوعی و یادگیری عمیق LSTM با یکدیگر بر مبنای شاخص‌های اعتبارسنجی RMSE، MAPE و AADT مقایسه می‌شوند.

در اعتبارسنجی به وسیله شاخص RMSE، بیشترین اختلاف خطا در مدل ARIMA و LSTM مربوط به مسیر چناران - مشهد با ۰,۷۲ می‌باشد. همچنین بیشترین اختلاف خطا در مدل شبکه عصبی مصنوعی و LSTM مربوط به مسیر مشهد - کلات با ۰,۱۸ می‌باشد.

در اعتبارسنجی به وسیله شاخص MAPE، بیشترین اختلاف خطا در مدل ARIMA و LSTM مربوط به مسیر ملک‌آباد - مشهد با ۲۴٪ می‌باشد. همچنین بیشترین اختلاف خطا در مدل شبکه عصبی مصنوعی و LSTM مربوط به مسیر مشهد - کلات با ۱۱٪ می‌باشد.

اعتبارسنجی به وسیله شاخص AADT، بیشترین اختلاف دقت پیش‌بینی در مدل ARIMA و LSTM مربوط به مسیر چناران - مشهد با ۲۳٪ اختلاف می‌باشد. در ضمن برای دو مسیر سدکارده - مشهد و مشهد - کلات روش ARIMA دقت پیش‌بینی بهتری نسبت به مدل LSTM دارد اما همان گونه توضیح داده شد مدل ARIMA اصلاً مدلسازی خوبی برای این مسیرها ارائه نداده است. همچنین بیشترین اختلاف دقت پیش‌بینی در مدل شبکه عصبی مصنوعی و LSTM مربوط به مسیر مشهد - کلات با ۶٪ می‌باشد. در ضمن در برخی از مسیرها مدل شبکه عصبی مصنوعی نسبت به LSTM با درصد اختلاف - های خیلی جزئی بهتر عمل کرده‌است.

یکی از مشکلات اصلی در پیش‌بینی جریان ترافیک ماهیت غیر خطی بودن داده‌ها می‌باشد. یکی از اصلی‌ترین دلایل غیرخطی بودن داده‌ها جریان ترافیک بستگی این داده‌ها به روزهای هفته از جمله پنجشنبه و جمعه‌ها با دیگر روزهای هفته می‌باشد. این عامل در جاده‌های توریستی - تفریحی نقش پررنگ تری دارند. همچنین می‌توان تعطیلی‌های چند روزه که در تقویم شمسی و قمری کشور وجود دارد را از عوامل اصلی غیرخطی بودن داده‌ها عنوان کرد.

از پژوهش حاضر دریافتیم که روش یادگیری عمیق LSTM در آموزش، مدلسازی و پیش‌بینی برای ۱۲ مسیر موجود عملکرد بهتری نسبت به روش‌های ARIMA و شبکه عصبی مصنوعی ارائه می‌دهد. مدل ARIMA بیشترین نوسانات و خطاهای پیش‌بینی را در این پژوهش دارد زیرا که این روش یک روش پارامتریک می‌باشد. روش‌های پارامتریک در داده‌هایی که خطی هستند و تغییرات منظمی دارند عملکرد خوبی را از خود نشان می‌دهند اما در داده‌هایی که ماهیت خیر خطی دارند عملکرد خوبی ندارند.

مدل شبکه عصبی مصنوعی یک روش غیر پارامتریک می‌باشد این روش همانطور که از نتایج پژوهش مشخص می‌باشد نتایج بهتری نسبت به مدل ARIMA دارد اما نمی‌توان این روش را یک روش کامل دانست چراکه این روش دارای معماری سطحی و در استخراج همبستگی داده‌ها عملکرد ضعیفی دارد. به طور کلی اکثر روش‌های غیرپارامتریک نمی‌توانند در اطلاعات عمیق ترافیکی نفوذ کنند. اما در مقابل این دو روش روش یادگیری عمیق LSTM قدرت این را دارد که از وابستگی بلندمدت داده‌ها جریان ترافیک استفاده کند و ویژگی‌های پنهان در جریان ترافیک را کشف کند و عملکرد پیش‌بینی بهتری را ارائه دهد.

۵ - ۴ پیشنهادات

- در سری‌های زمانی تشخیص ماتریس همبستگی از مهمترین بخش‌های مدلسازی و پیش‌بینی می‌باشد. تشخیص‌های دیگر در مورد ماتریس همبستگی در مدلسازی و پیش‌بینی مدل بررسی گردد.
- استفاده از روش‌های ترکیبی برای مدلسازی و پیش‌بینی جریان ترافیک و مقایسه هر کدام از روش‌ها
- در پژوهش حاضر داده‌های حجم تردد به صورت خام در مدلسازی استفاده گردیده‌اند. از اصلاح و خوشه بندی داده‌های ورودی استفاده گردد و تاثیر آن بر مدل بررسی شود.
- شرایط جوی آب و هوا یکی از شرایط تاثیرگذار بر حجم تردد می‌باشد استفاده از شرایط جوی و تاثیر این پارامتر در مدلسازی و پیش‌بینی بررسی گردد.
- استفاده عملی و تئوری از روش‌های نوین مدلسازی و پیش‌بینی در مدیریت حمل و نقل و ترافیک

- استفاده از نتایج پژوهش جهت پیش بینی حجم تردد با صرف هزینه و زمان حداقل برای جمع‌آوری داده‌ها
- تفکیک مسیرهای مورد ارزیابی از نظر حجم تردد، نوع جاده، کاربرد جاده و ...

مراج

اسدی فارسانی پ، (۱۳۹۴)، پایان نامه ارشد: "استفاده از شبکه عصبی برای پیش‌بینی کوتاه مدت جریان ترافیک آزادراه‌ها"، دانشکده عمران، دانشگاه یزد.

اسماعیلیان، م.، ربیعی، م. (۱۳۹۴) "راهنمای جامع SPSS۲۲" جلد اول، چاپ اول، موسسه فرهنگی هنری دیباگران تهران، تهران.

انصاری ص، (۱۳۹۳)، پایان نامه ارشد: "طراحی شبکه عصبی پویا به منظور پیش‌بینی حجم ترافیک"، دانشکده کامپیوتر، دانشگاه خواجه نصیرالدین طوسی.

جهانگرد، م.، پیرایش، م.ع. و محمدزاده مقدم، ا.ا.، (۱۳۹۵)، "پیش‌بینی زمان سفر در شبکه با استفاده از سری زمانی (مطالعه موردی: ناحیه ۱ مشهد)" فصلنامه جاده، دوره ۲۴، شماره ۸۹، زمستان ۱۳۹۵، ص. ۲۱۸-۲۰۹.

ماکریداکیس، ا.، ویل رایت، س.ک.، هایندمن، ر.ج. (۱۳۹۴) "پیش‌بینی، روش‌ها و کاربردها" جلد اول، شاهسونی، د.آرشی، م. چاپ اول، دانشگاه صنعتی شاهرود، شاهرود.

معتمدی سده م، (۱۳۹۳)، پایان نامه ارشد: "ارایه مدل پیش‌بینی حجم روزانه ترافیک برون شهری (ADT) با استفاده از شبکه‌های عصبی"، دانشکده عمران، دانشگاه پیام نور.

Ahmed, M. S., & Cook, A. R. (۱۹۷۹). Analysis of freeway traffic time-series data by using Box-Jenkins techniques (No. ۷۲۲).

Bernaś, M., Placzek, B., Porwik, P., & Pamuła, T. (۲۰۱۵). Segmentation of vehicle detector data for improved k-nearest neighbours-based traffic flow prediction. *IET intelligent transport systems*, ۹(۳), ۲۶۴-۲۷۴.

Castro-Neto, M., Jeong, Y. S., Jeong, M. K., & Han, L. D. (۲۰۰۹). Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions. *Expert systems with applications*, ۳۶(۳), ۶۱۶۴-۶۱۷۳.

Chen, C., Hu, J., Meng, Q. and Zhang, Y., (۲۰۱۱), "Short-time traffic flow prediction with ARIMA-GARCH model," ۲۰۱۱ *IEEE Intelligent Vehicles Symposium (IV)*, pp ۶۰۷-۶۱۲.

Chen, D. (۲۰۱۷). Research on traffic flow prediction in the big data environment based on the improved RBF neural network. *IEEE Transactions on Industrial Informatics*, ۱۳(۴), ۲۰۰۰-۲۰۰۸.

- Dogan, E. (2020). Analysis and comparison of long short-term memory networks short-term traffic prediction performance. *Zeszyty Naukowe. Transport/Politechnika Śląska*.
- Farokhi Sadabadi, K., Hamed, M., & Haghani, A. (2010). Evaluating moving average techniques in short-term travel time prediction using an AVI data set (No. 10-3144).
- Fei, X., Lu, C. C., & Liu, K. (2011). A bayesian dynamic linear model approach for real-time short-term freeway travel time prediction. *Transportation Research Part C: Emerging Technologies*, 19(6), 1306-1318.
- Ghosh, B., Basu, B. and O'Mahony, M., (2005), "Time-series modelling for forecasting vehicular traffic flow in Dublin," *In 14th Annual Meeting of the Transportation Research Board*, Washington, DC.
- Guin, A., (2006), "Travel Time Prediction using a Seasonal Autoregressive Integrated Moving Average Time Series Model," 2006 *IEEE Intelligent Transportation Systems Conference*, pp 493-498.
- Guo, J., Huang, W., & Williams, B. M. (2014). Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification. *Transportation Research Part C: Emerging Technologies*, 33, 50-64.
- Jia, Y., Wu, J., & Xu, M. (2017). Traffic flow prediction with rainfall impact using a deep learning method. *Journal of advanced transportation*, 2017.
- Khan, Z., Khan, S. M., Dey, K., & Chowdhury, M. (2019). Development and evaluation of recurrent neural network-based models for hourly traffic volume and annual average daily traffic prediction. *Transportation Research Record*, 2673(7), 489-503.
- Levin, M., & Tsao, Y. D. (1980). On forecasting freeway occupancies and volumes (abridgment). *Transportation Research Record*, (773).
- Li, L., Chen, X., Li, Z., & Zhang, L. (2013). Freeway travel-time estimation based on temporal-spatial queueing model. *IEEE Transactions on Intelligent Transportation Systems*, 14(3), 1536-1541.
- Luo, X., Li, D., Yang, Y., & Zhang, S. (2019). Spatiotemporal traffic flow prediction with KNN and LSTM. *Journal of Advanced Transportation*, 2019.
- Ma, X., Tao, Z., Wang, Y., Yu, H., & Wang, Y. (2015). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C: Emerging Technologies*, 54, 187-197.

Shao, H., & Soong, B. H. (2016, November). Traffic flow prediction with long short-term memory networks (LSTMs). In 2016 IEEE region 10 conference (TENCON) (pp. 2986-2989). IEEE.

Tian, Y., & Pan, L. (2015, December). Predicting short-term traffic flow by long short-term memory recurrent neural network. In 2015 IEEE international conference on smart city/SocialCom/SustainCom (SmartCity) (pp. 153-158). IEEE.

Van Der Voort, M., Dougherty, M., & Watson, S. (1996). Combining Kohonen maps with ARIMA time series models to forecast traffic flow. *Transportation Research Part C: Emerging Technologies*, 4(4), 307-318.

Wang, J., Deng, W., & Guo, Y. (2014). New Bayesian combination method for short-term traffic flow forecasting. *Transportation Research Part C: Emerging Technologies*, 32, 79-94.

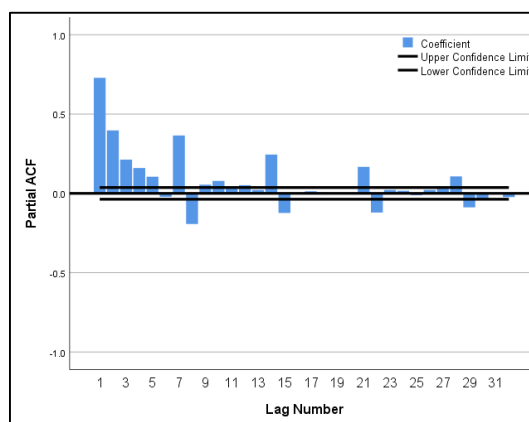
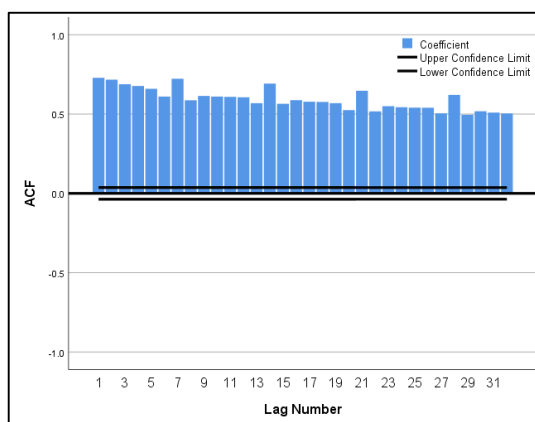
Wei, Y., & Chen, M. C. (2012). Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks. *Transportation Research Part C: Emerging Technologies*, 21(1), 148-162.

Williams, B. M., & Hoel, L. A. (2003). Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results. *Journal of transportation engineering*, 129(6), 664-672.

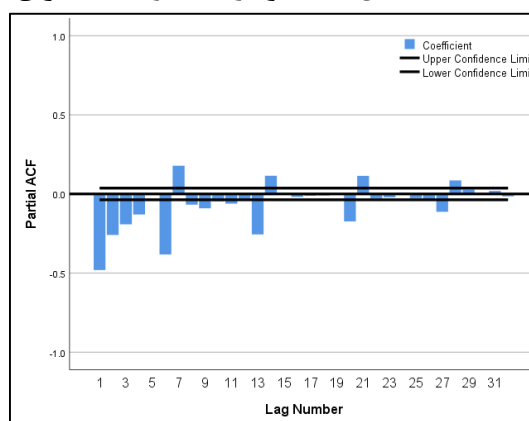
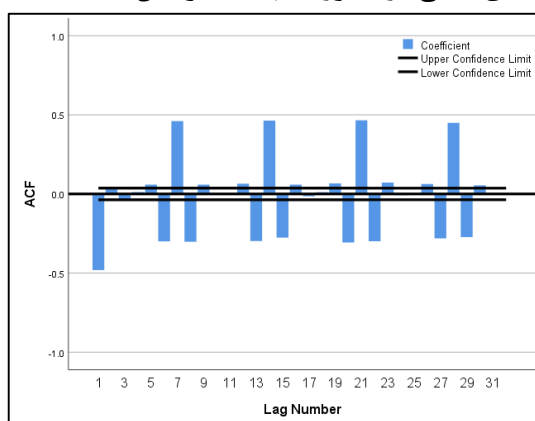
Wu, J., Zhong, L., Li, L. and Lu, A., (2013), "A prediction model based on time series data in Intelligent Transportation System," In *International Conference on Information Computing and Applications*, pp 420-429.

Zhao, Z., Chen, W., Wu, X., Chen, P. C., & Liu, J. (2017). LSTM network: a deep learning approach for short-term traffic forecast. *IET Intelligent Transport Systems*, 11(2), 68-75.

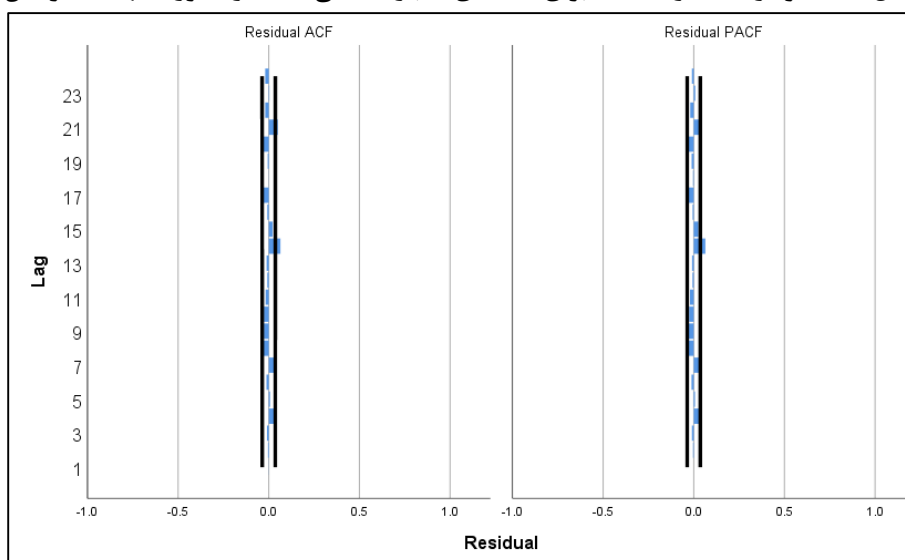
سوت



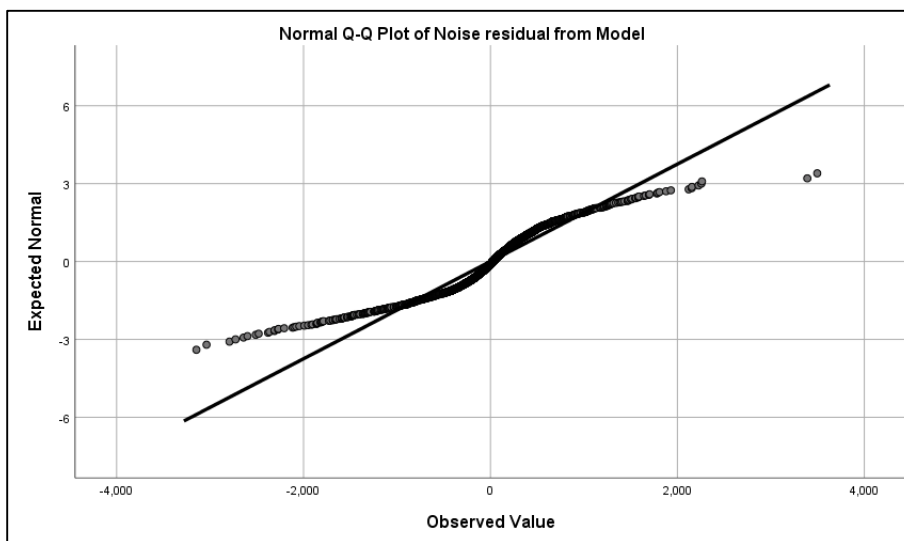
شکل آ- ۱ نمودار ACF و PACF برای داده‌های اصلی در محور مشهد - سرخس



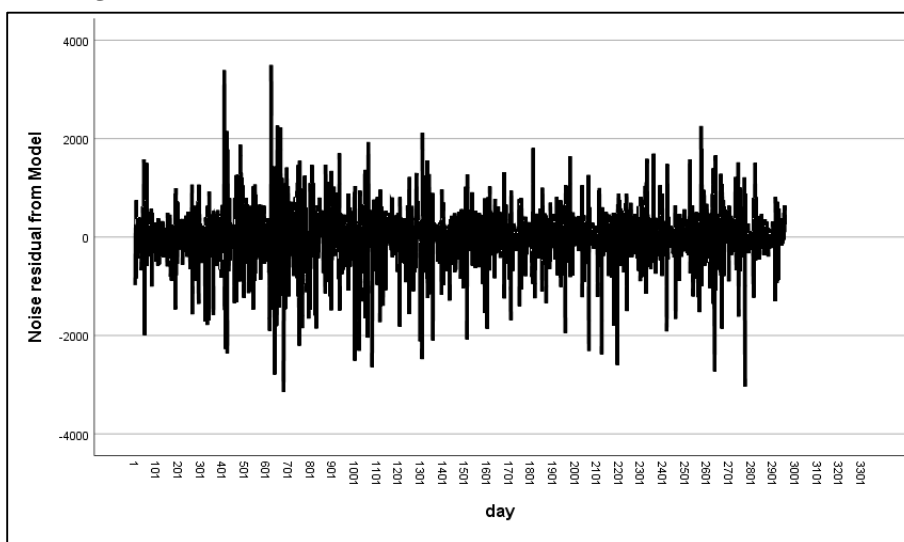
شکل آ- ۲ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور مشهد - سرخس



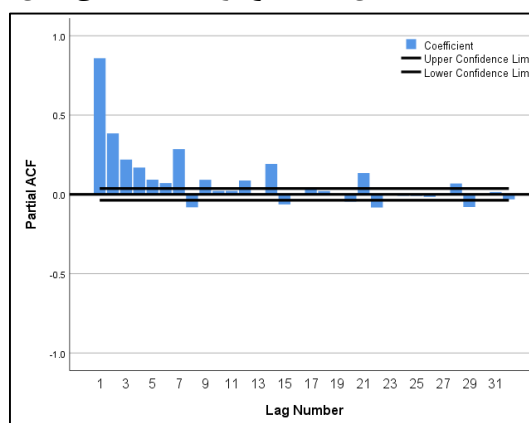
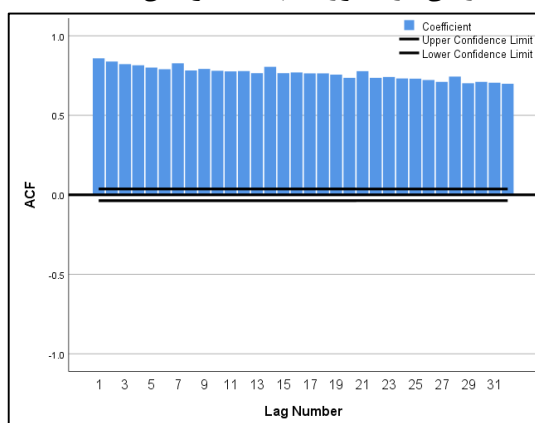
شکل آ- ۳ نمودارهای ACF و PACF باقیمانده‌های در محور مشهد - سرخس



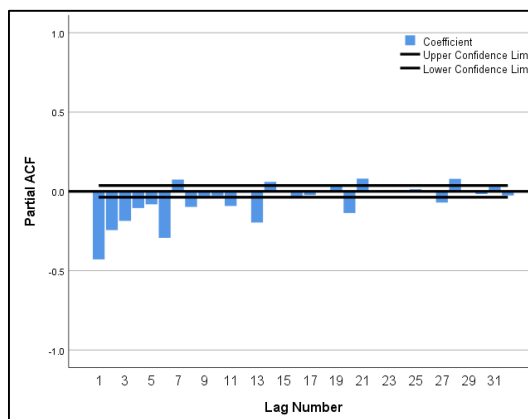
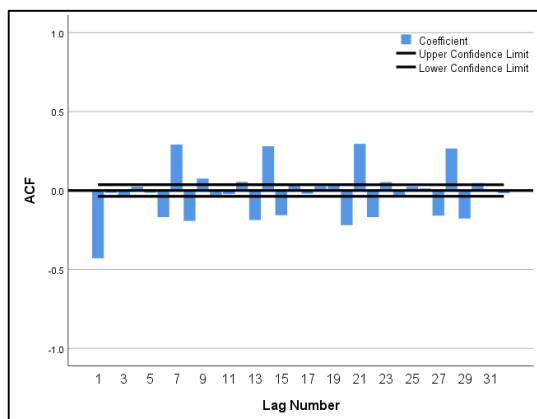
شکل آ- ۴ نمودار احتمال نرمال باقیمانده‌ها در محور مشهد - سرخس



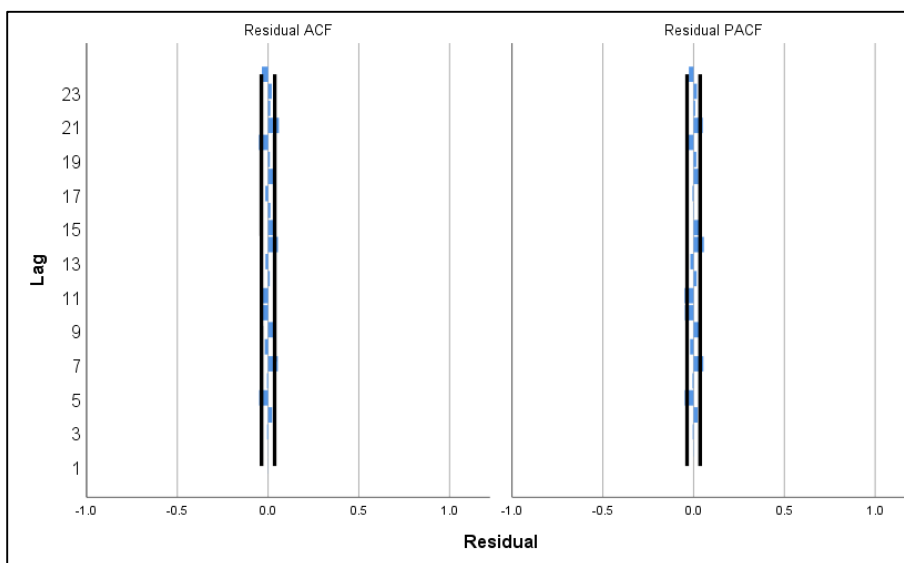
شکل آ- ۵ نمودار باقیمانده‌های مدل نسبت به زمان در محور مشهد - سرخس



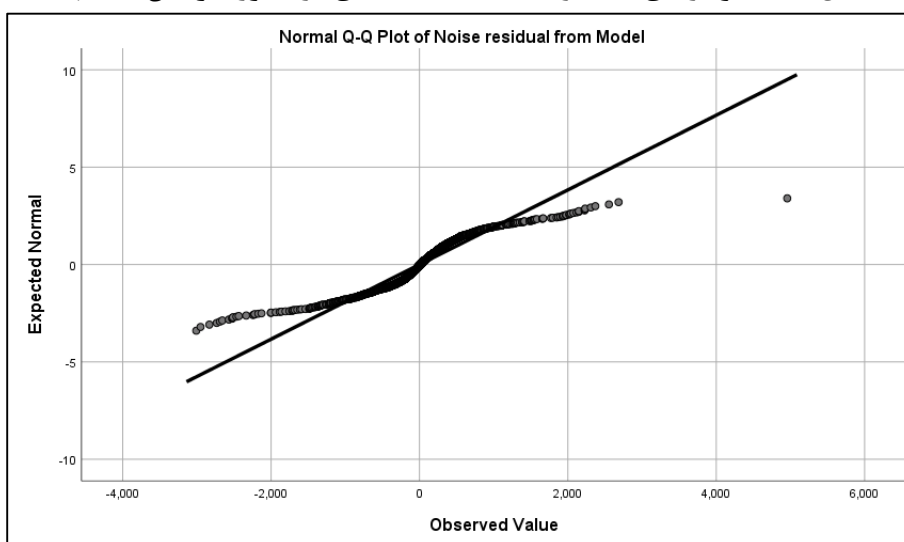
شکل آ- ۶ نمودار ACF و PACF برای داده‌های اصلی در محور مشهد - سرخس



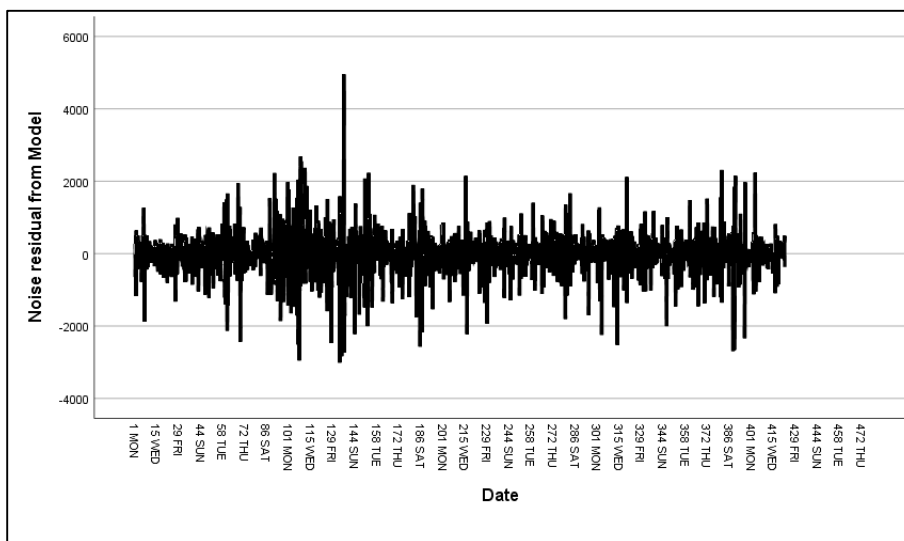
شکل آ- ۷ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور سرخس - مشهد



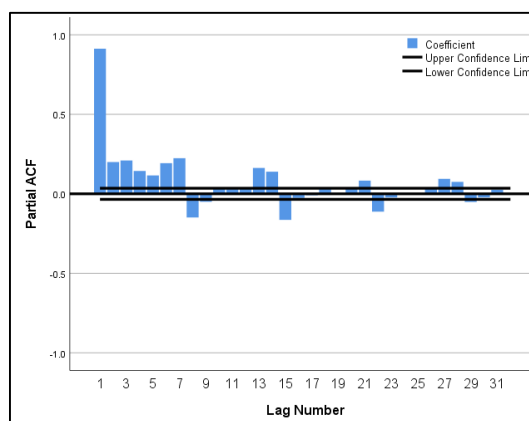
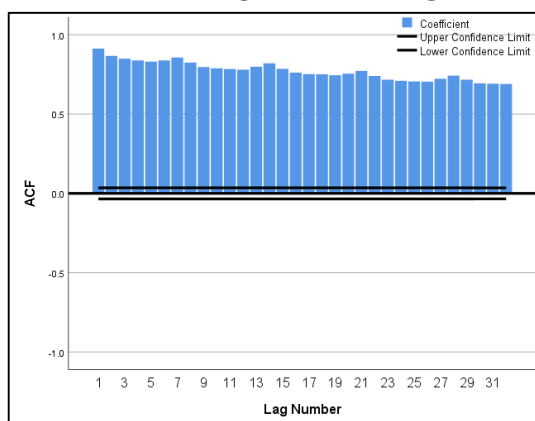
شکل آ- ۸ نمودارهای ACF و PACF باقیمانده‌های در محور سرخس - مشهد



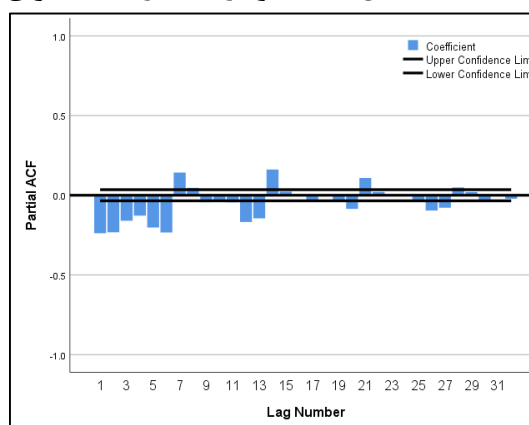
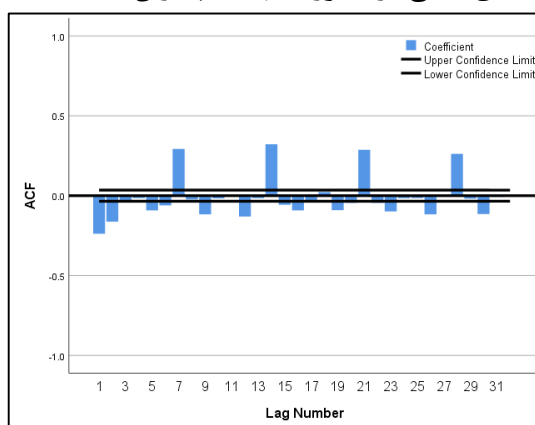
شکل آ- ۹ نمودار احتمال نرمال باقیمانده‌ها در محور سرخس - مشهد



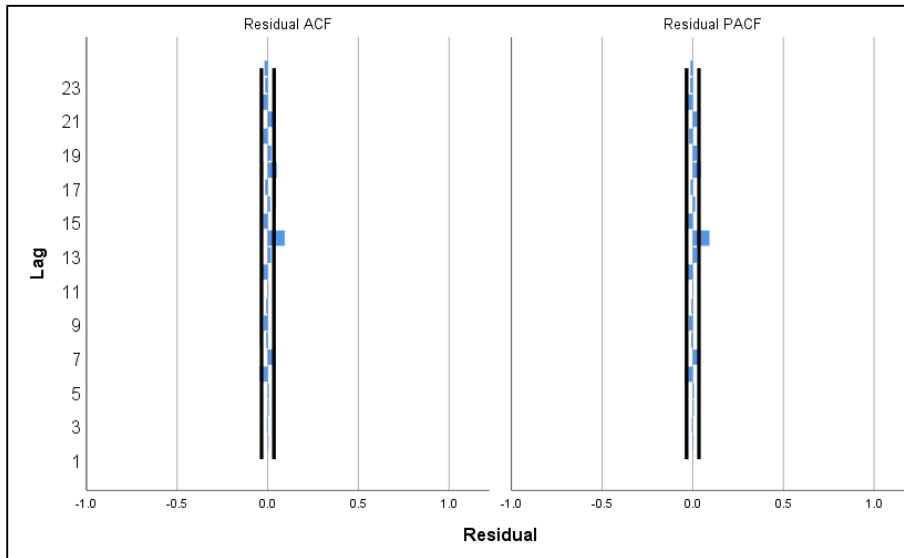
شکل آ- ۱۰ نمودار باقیمانده‌های مدل نسبت به زمان در محور سرخس - مشهد



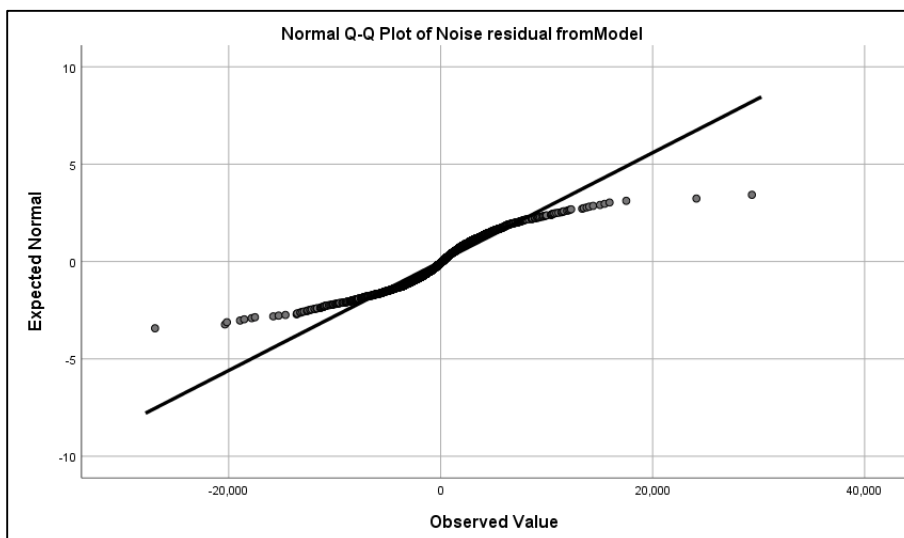
شکل آ- ۱۱ نمودار ACF و PACF برای داده‌های اصلی در محور مشهد - چناران



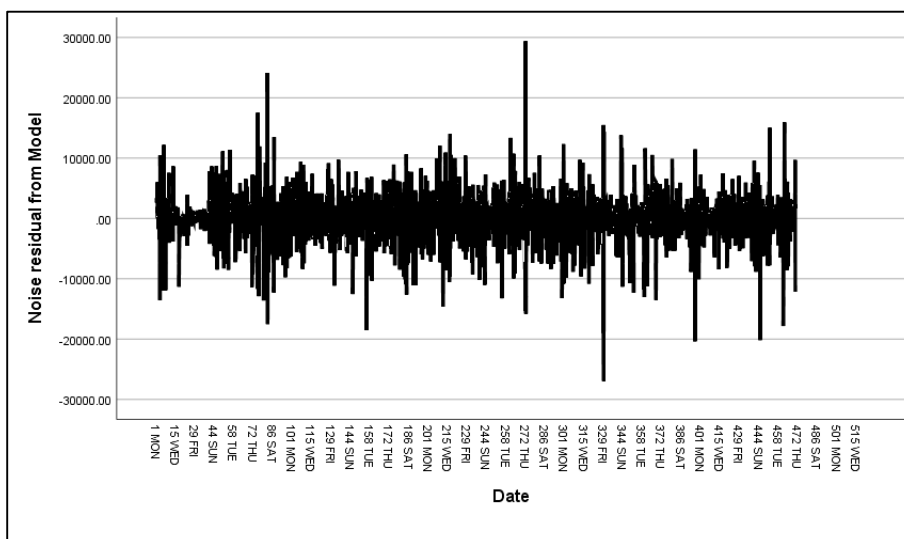
شکل آ- ۱۲ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور مشهد - چناران



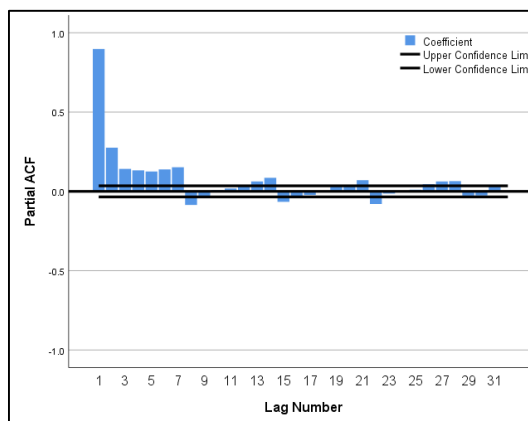
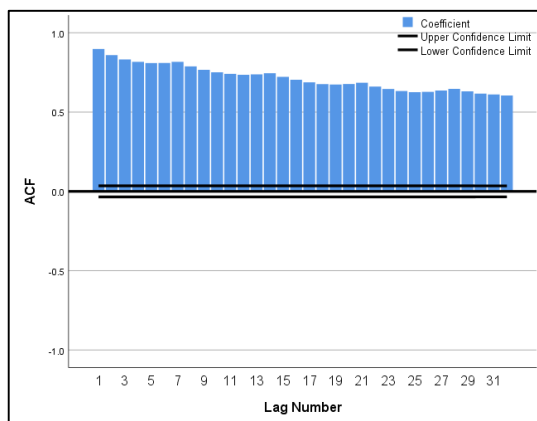
شکل آ- ۱۳ نمودارهای ACF و PACF باقیمانده‌های در محور مشهد - چناران



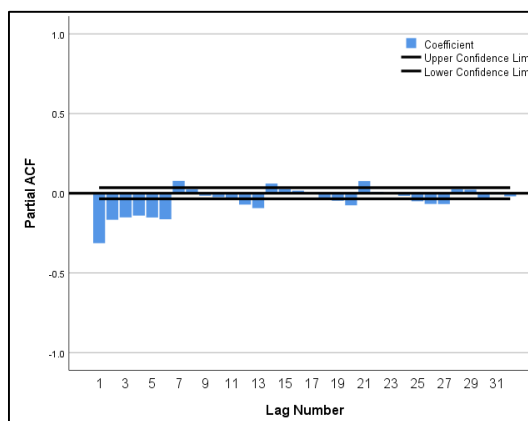
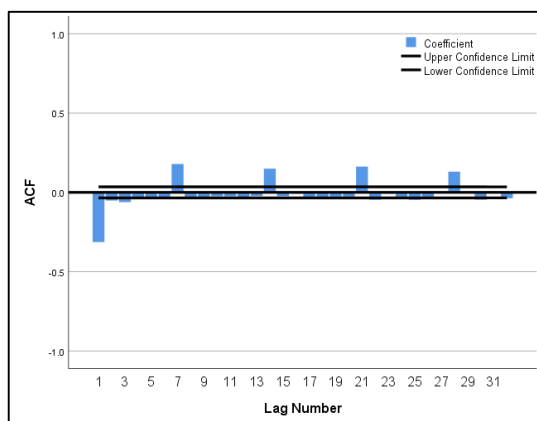
شکل آ- ۱۴ نمودار احتمال نرمال باقیمانده‌ها در محور مشهد - چناران



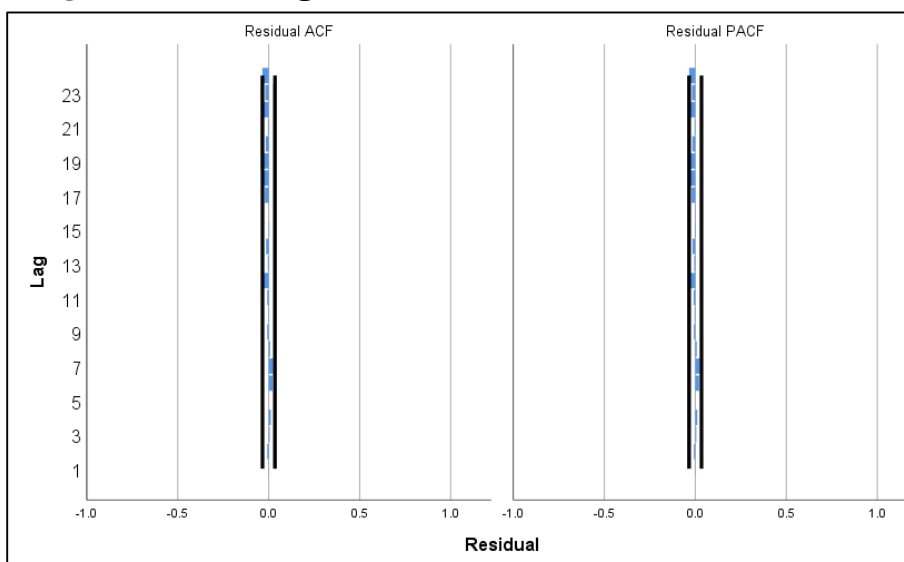
شکل آ- ۱۵ نمودار باقیمانده‌های مدل نسبت به زمان در محور مشهد - چناران



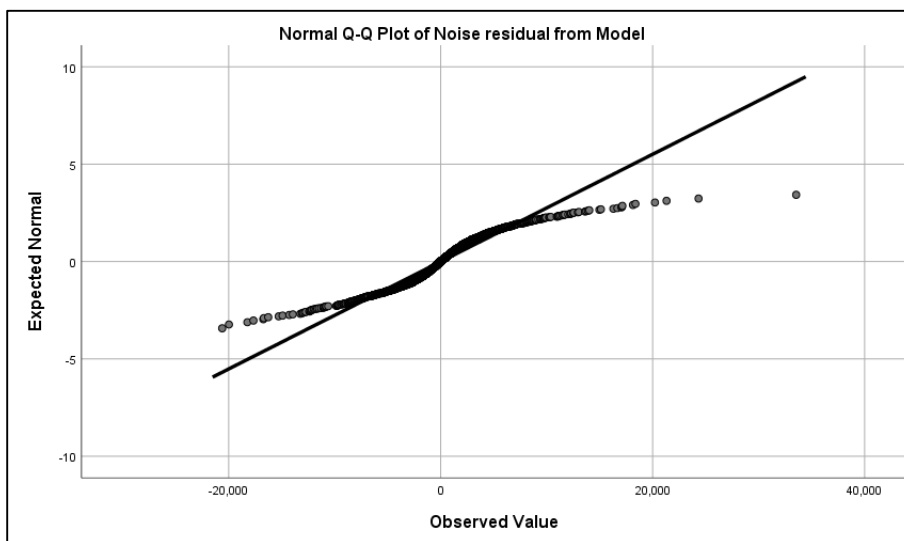
شکل آ- ۱۶ نمودار ACF و PACF برای داده‌های اصلی در محور چناران - مشهد



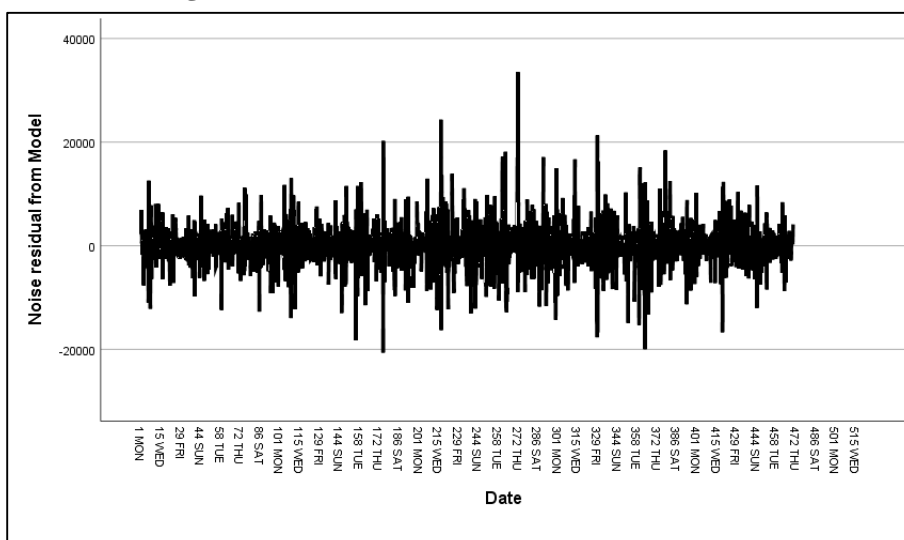
شکل آ- ۱۷ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور چناران - مشهد



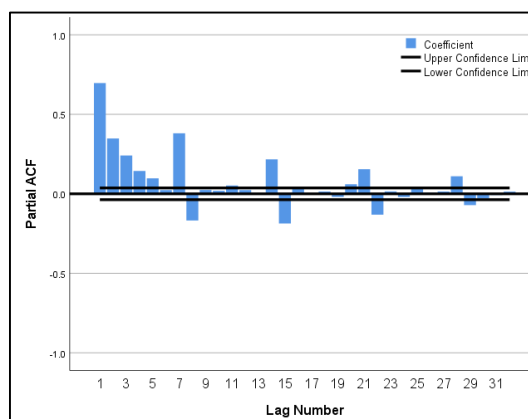
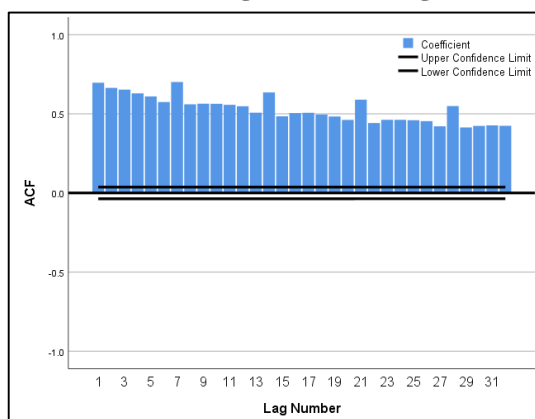
شکل آ- ۱۸ نمودارهای ACF و PACF باقیمانده‌های در محور چناران - مشهد



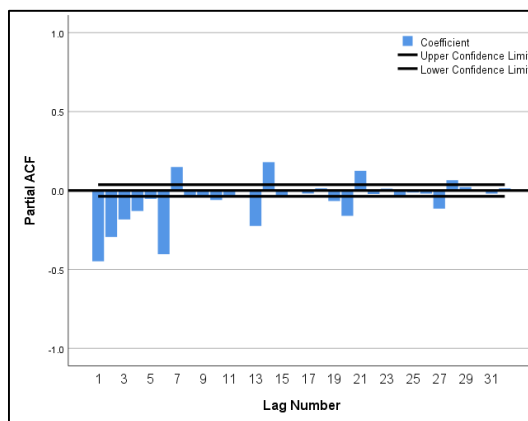
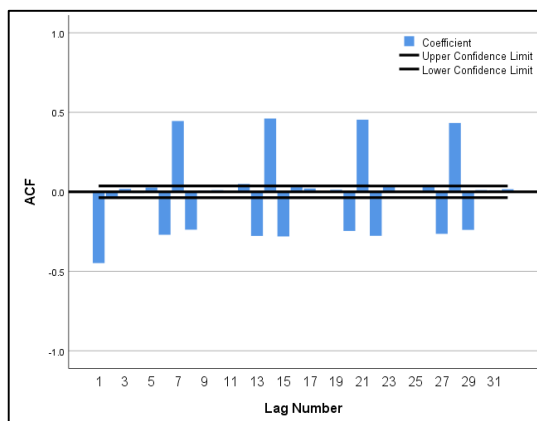
شکل آ- ۱۹ نمودار احتمال نرمال باقیمانده‌ها در محور چناران - مشهد



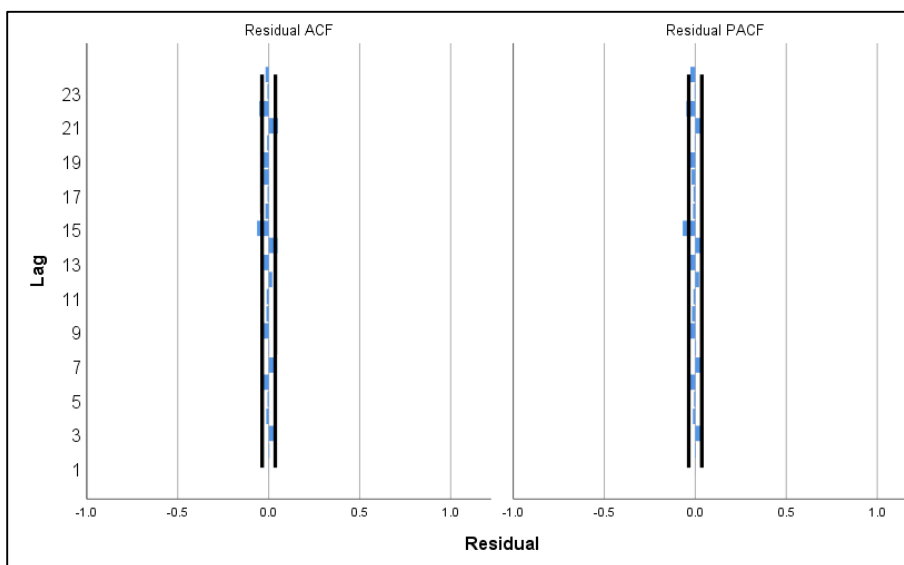
شکل آ- ۲۰ نمودار باقیمانده‌های مدل نسبت به زمان در محور چناران - مشهد



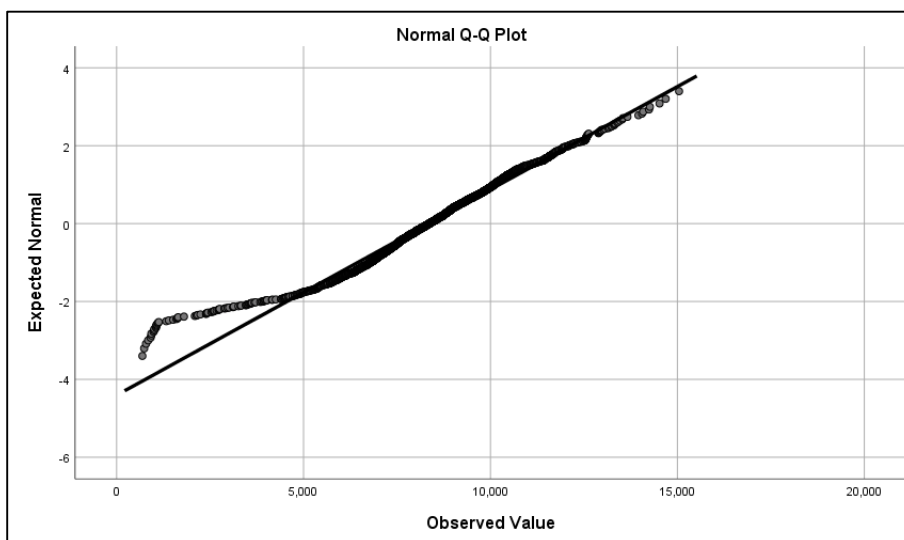
شکل آ- ۲۱ نمودار ACF و PACF برای داده‌های اصلی در محور مشهد - فریمان



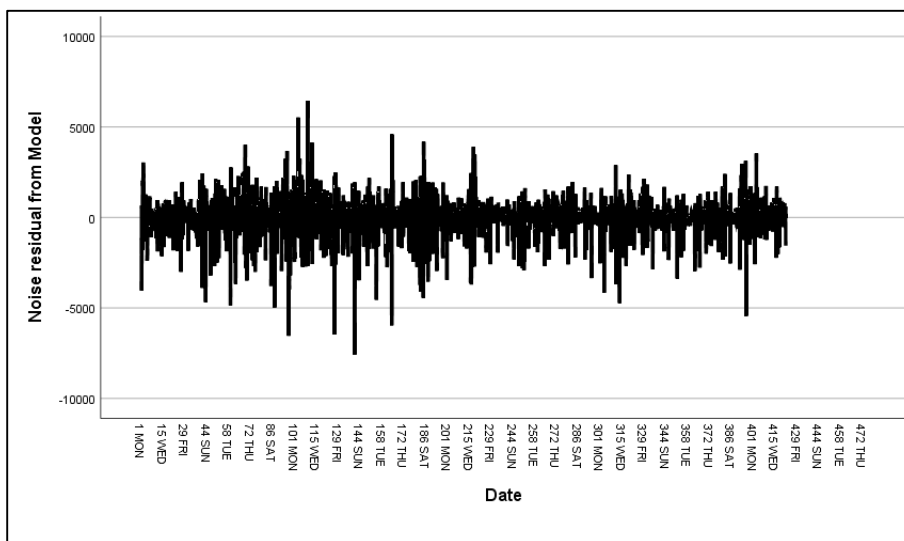
شکل آ- ۲۲ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور مشهد - فریمان



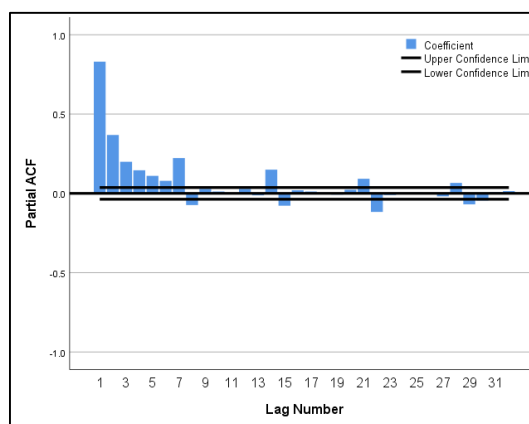
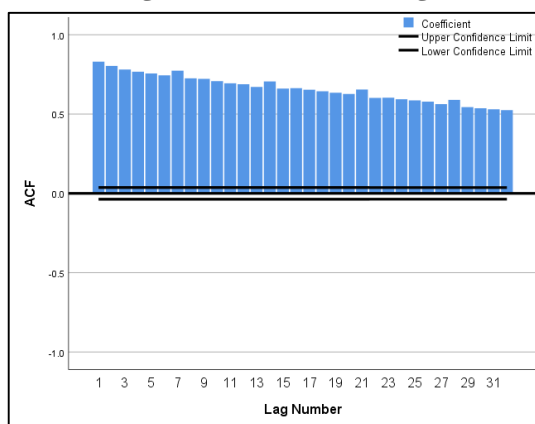
شکل آ- ۲۳ نمودارهای ACF و PACF باقیمانده‌های در محور مشهد - فریمان



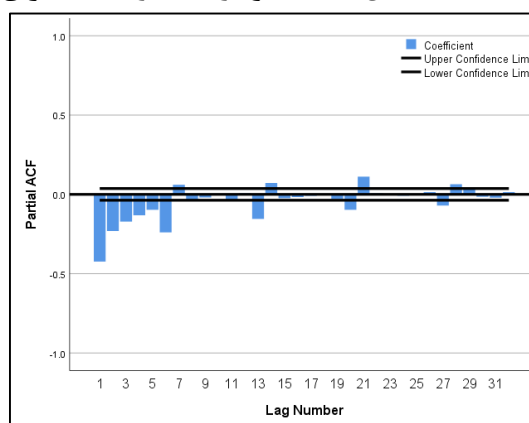
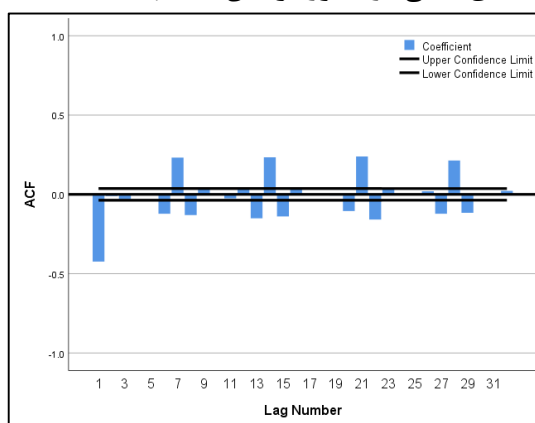
شکل آ- ۲۴ نمودار احتمال نرمال باقیمانده‌ها در محور مشهد - فریمان



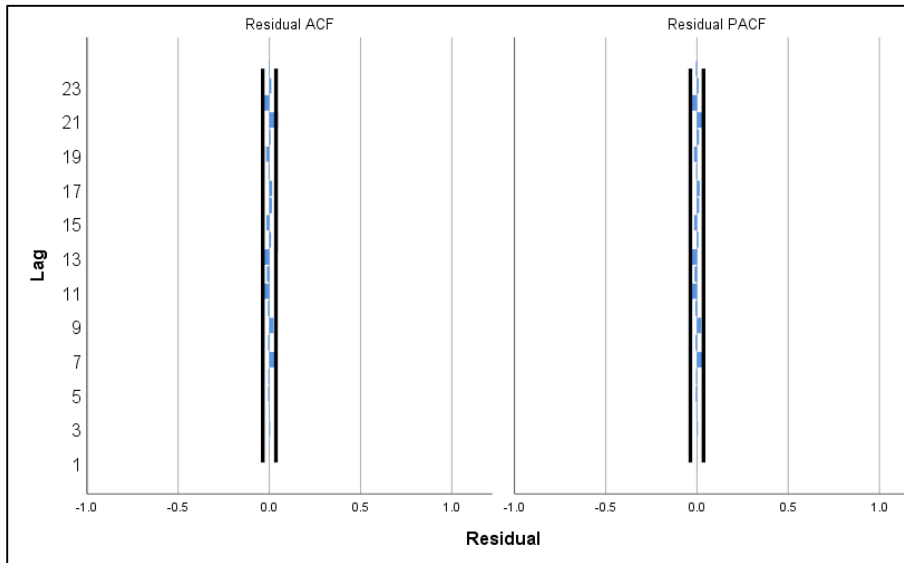
شکل آ- ۲۵ نمودار باقیمانده‌های مدل نسبت به زمان در محور مشهد - فریمان



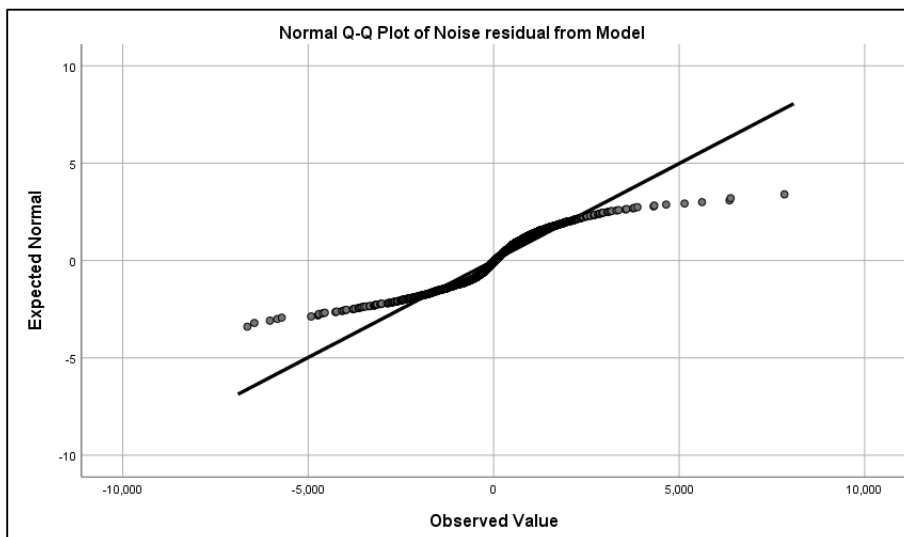
شکل آ- ۲۶ نمودار ACF و PACF برای داده‌های اصلی در محور فریمان - مشهد



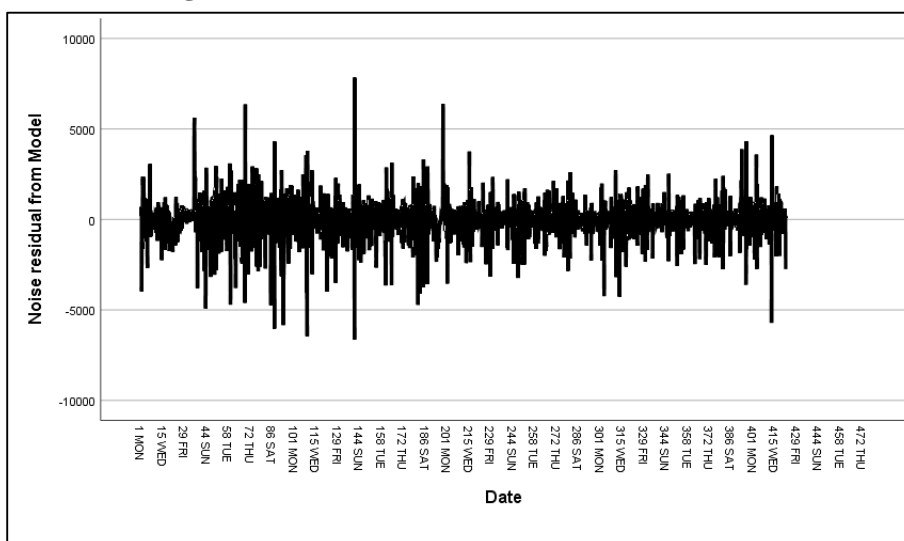
شکل آ- ۲۷ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور فریمان - مشهد



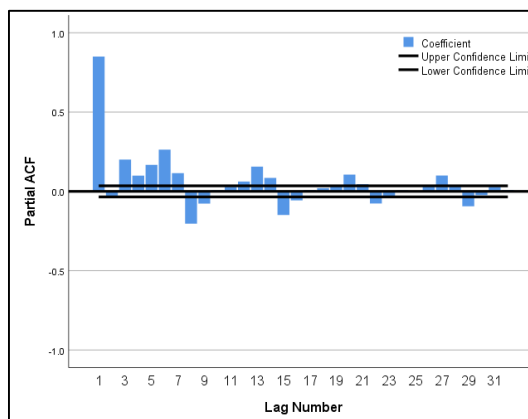
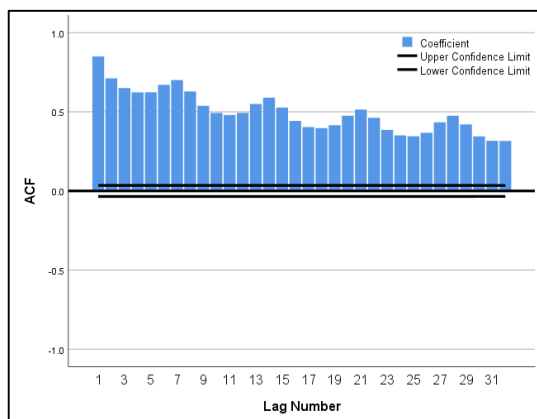
شکل آ- ۲۸ نمودارهای ACF و PACF باقیمانده‌های در محور فریمان - مشهد



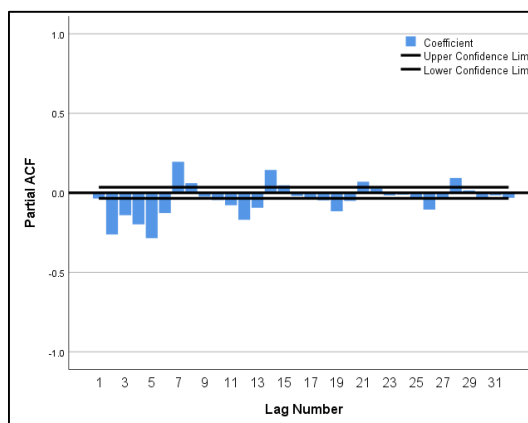
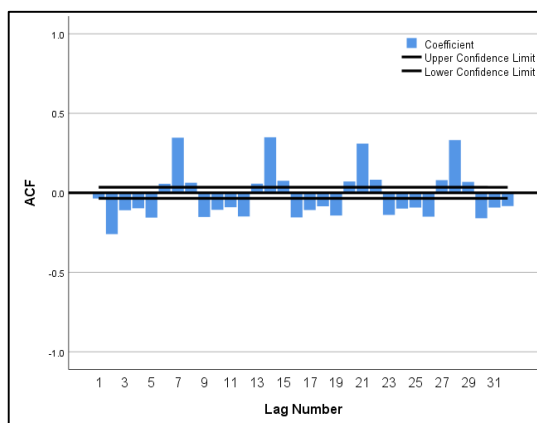
شکل آ- ۲۹ نمودار احتمال نرمال باقیمانده‌ها در محور فریمان - مشهد



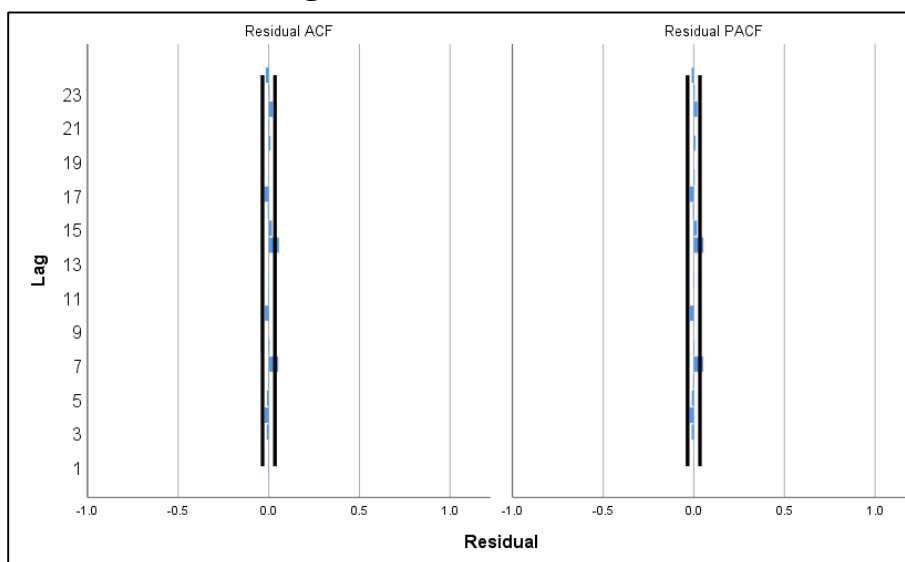
شکل آ- ۳۰ نمودار باقیمانده‌های مدل نسبت به زمان در محور فریمان - مشهد



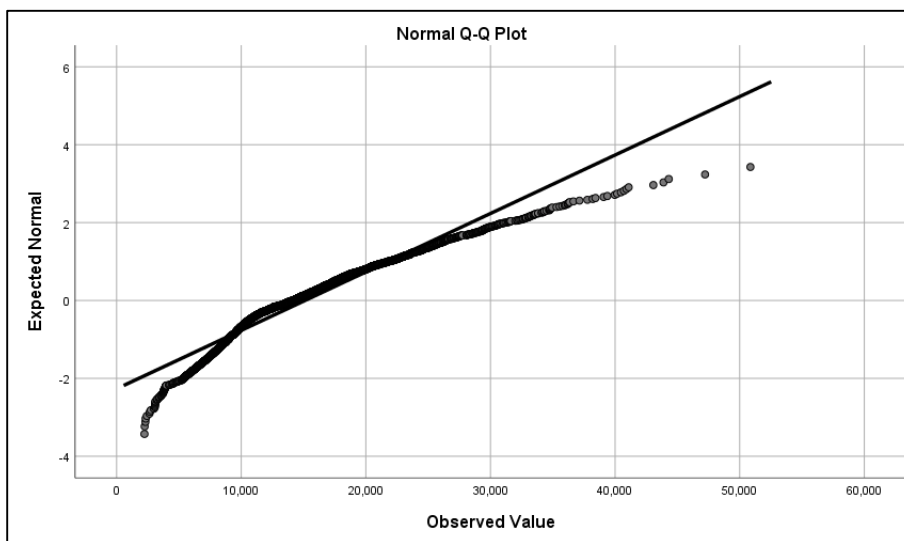
شکل آ- ۳۱ نمودار ACF و PACF برای داده‌های اصلی در محور مشهد - ملک آباد



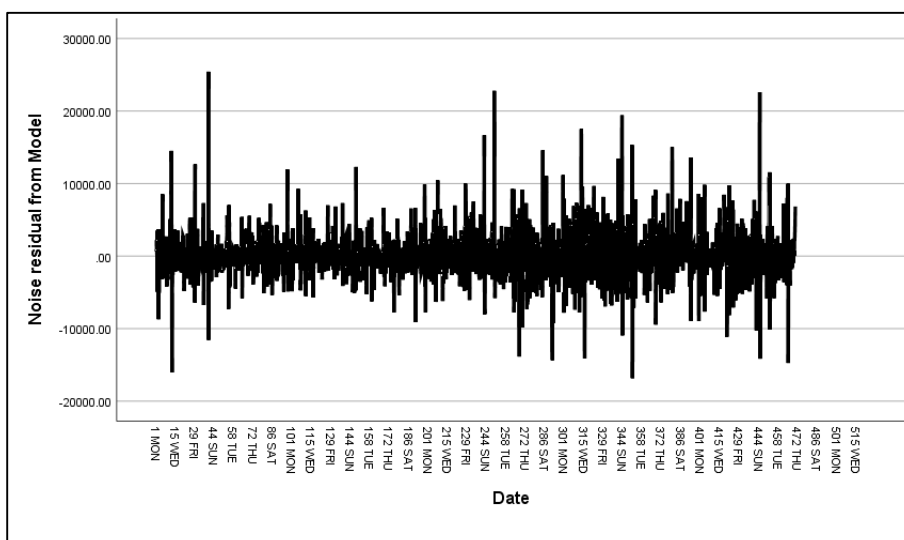
شکل آ- ۳۲ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور مشهد - ملک آباد



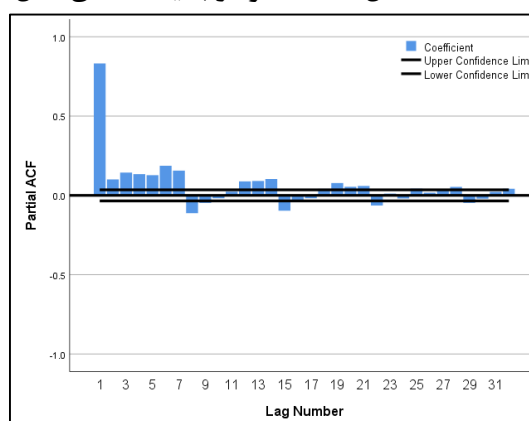
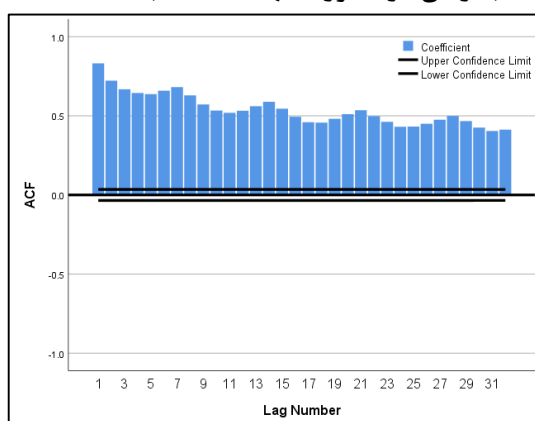
شکل آ- ۳۳ نمودارهای ACF و PACF باقیمانده‌های در محور مشهد - ملک آباد



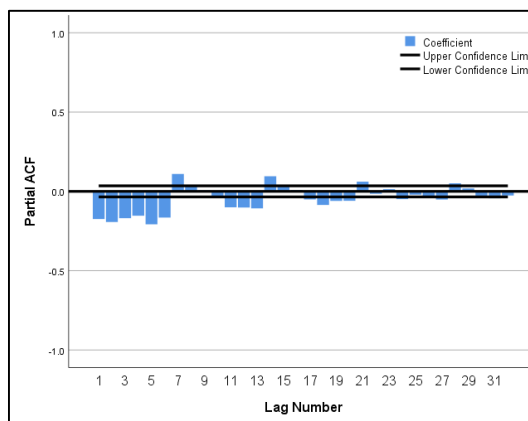
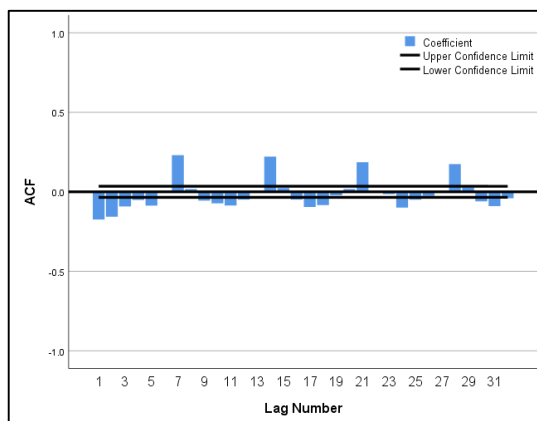
شکل آ- ۳۴ نمودار احتمال نرمال باقیمانده‌ها در محور مشهد - ملک آباد



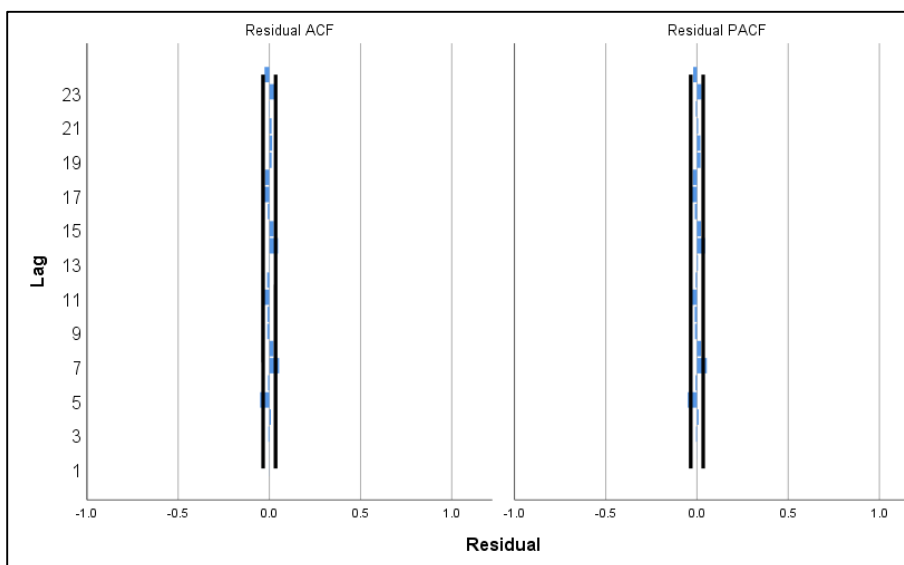
شکل آ- ۳۵ نمودار باقیمانده‌های مدل نسبت به زمان در محور مشهد - ملک آباد



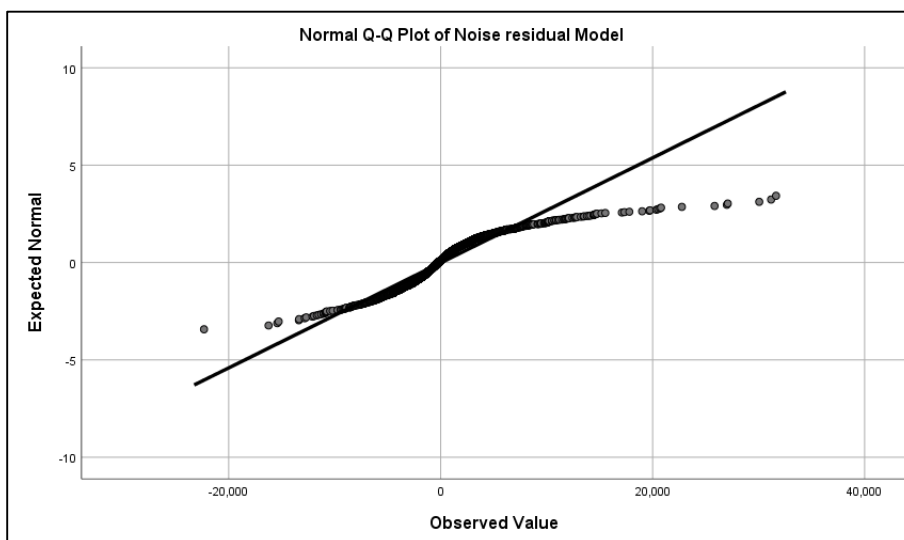
شکل آ- ۳۶ نمودار ACF و PACF برای داده‌های اصلی در محور ملک آباد - مشهد



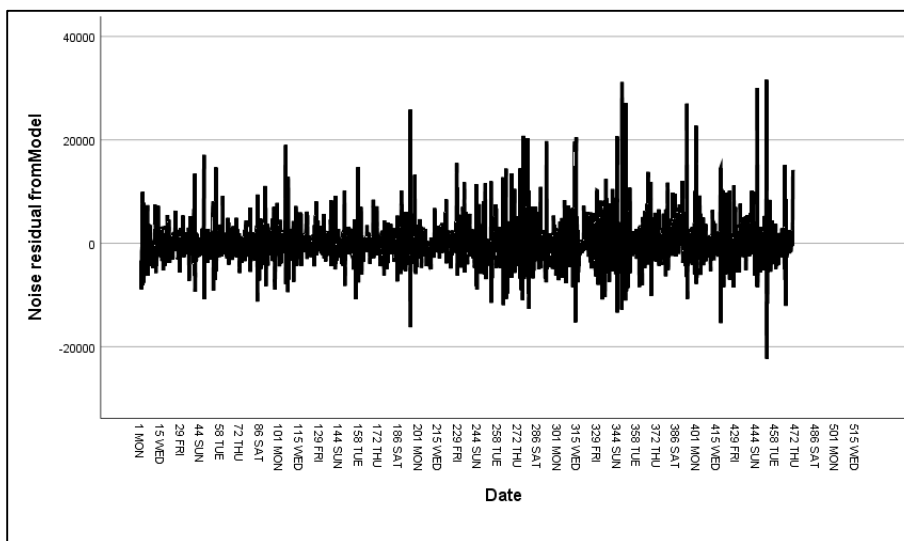
شکل آ- ۳۷ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور ملک آباد - مشهد



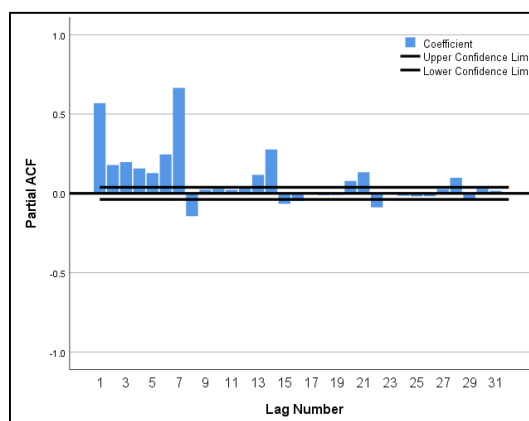
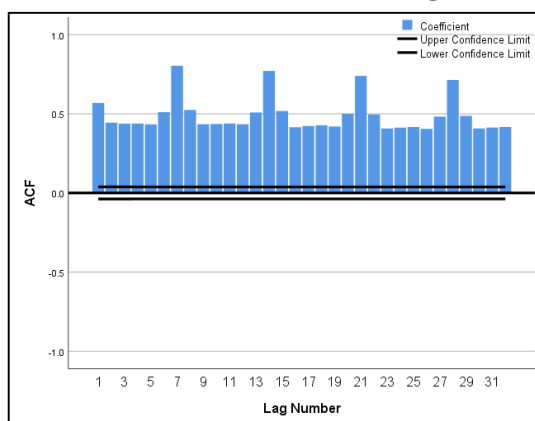
شکل آ- ۳۸ نمودارهای ACF و PACF باقیمانده‌های در محور ملک آباد - مشهد



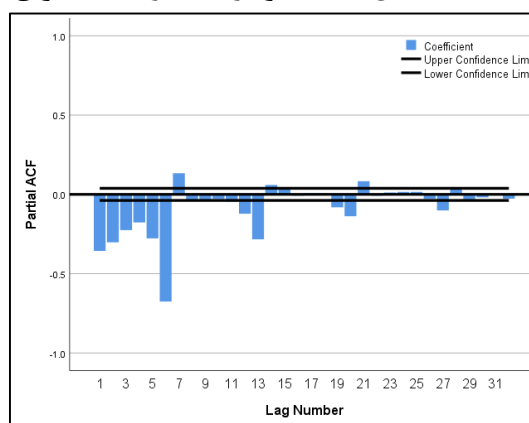
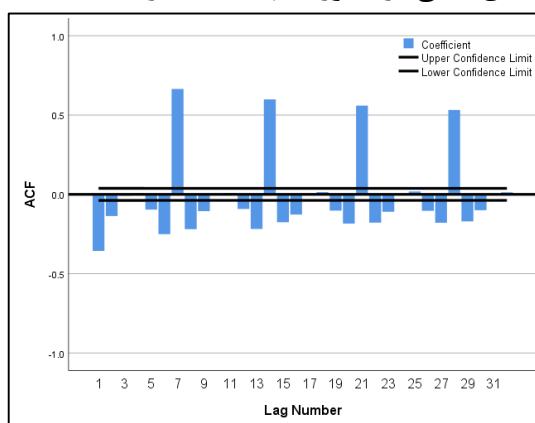
شکل آ- ۳۹ نمودار احتمال نرمال باقیمانده‌ها در محور ملک آباد - مشهد



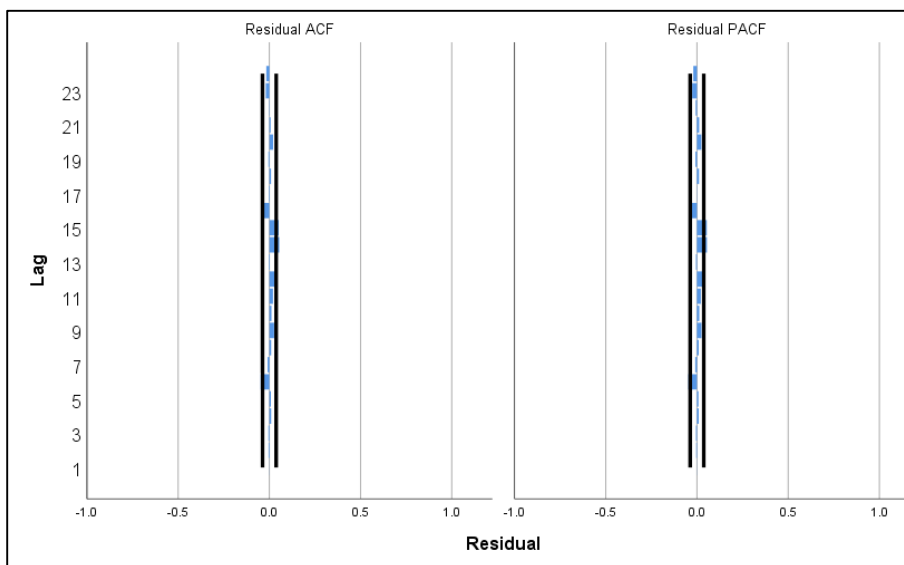
شکل آ- ۴۰ نمودار باقیمانده‌های مدل نسبت به زمان در محور ملک آباد - مشهد



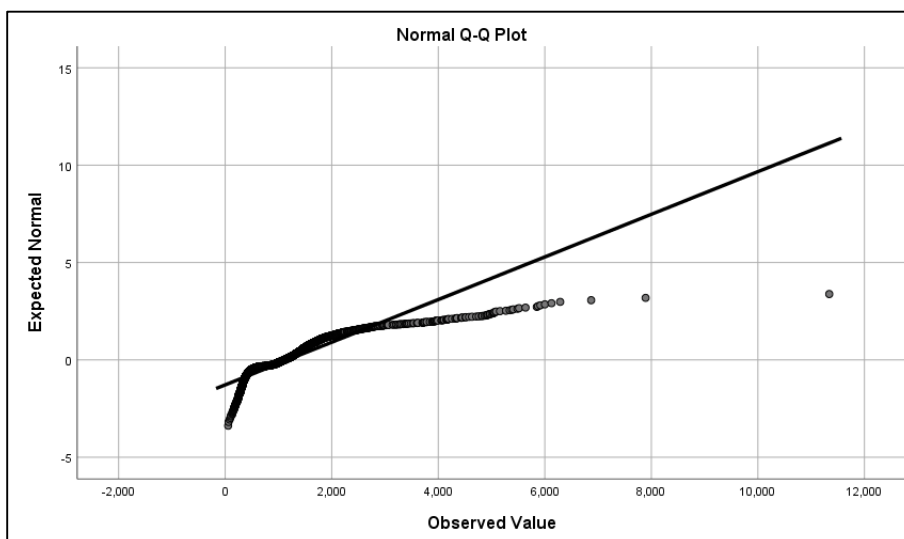
شکل آ- ۴۱ نمودار ACF و PACF برای داده‌های اصلی در محور مشهد - سدکارده



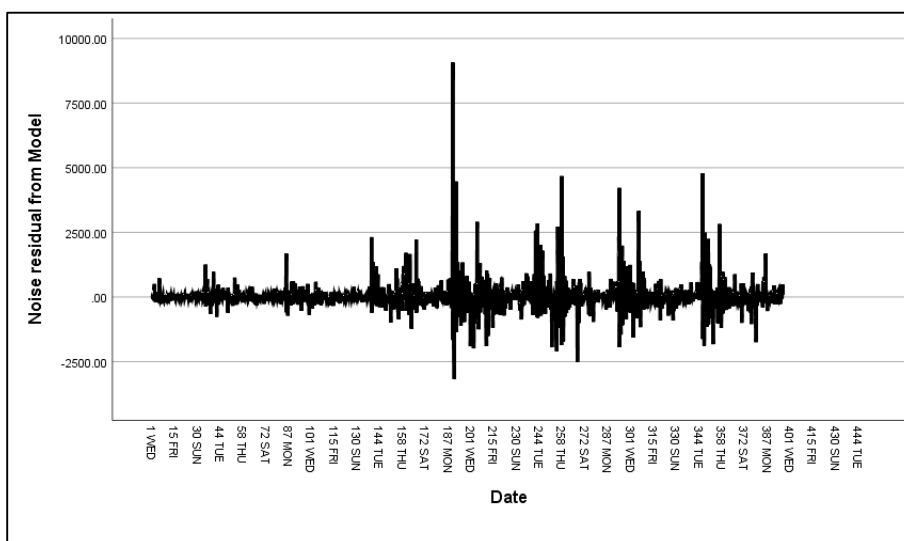
شکل آ- ۴۲ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور مشهد - سدکارده



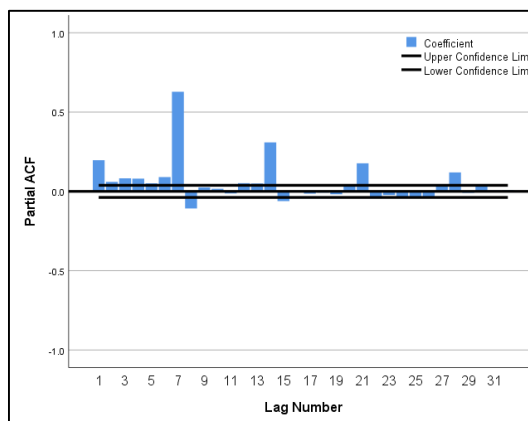
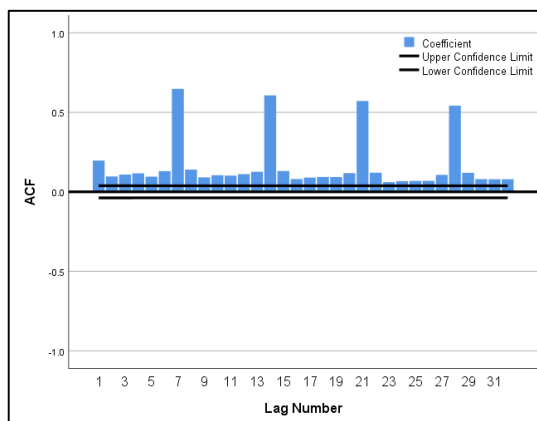
شکل آ- ۴۳ نمودارهای ACF و PACF باقیمانده‌های در محور مشهد - سدکارده



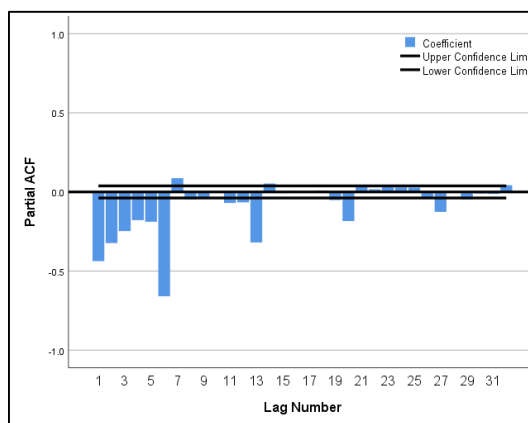
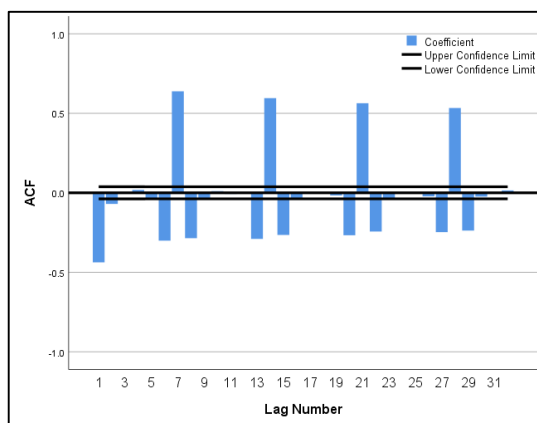
شکل آ- ۴۴ نمودار احتمال نرمال باقیمانده‌ها در محور مشهد - سدکارده



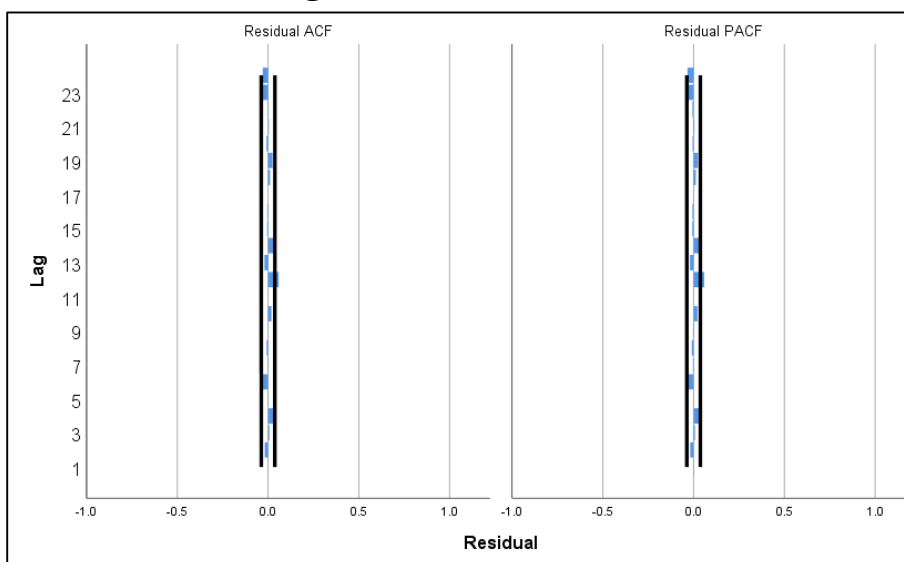
شکل آ- ۴۵ نمودار باقیمانده‌های مدل نسبت به زمان در محور مشهد - سدکارده



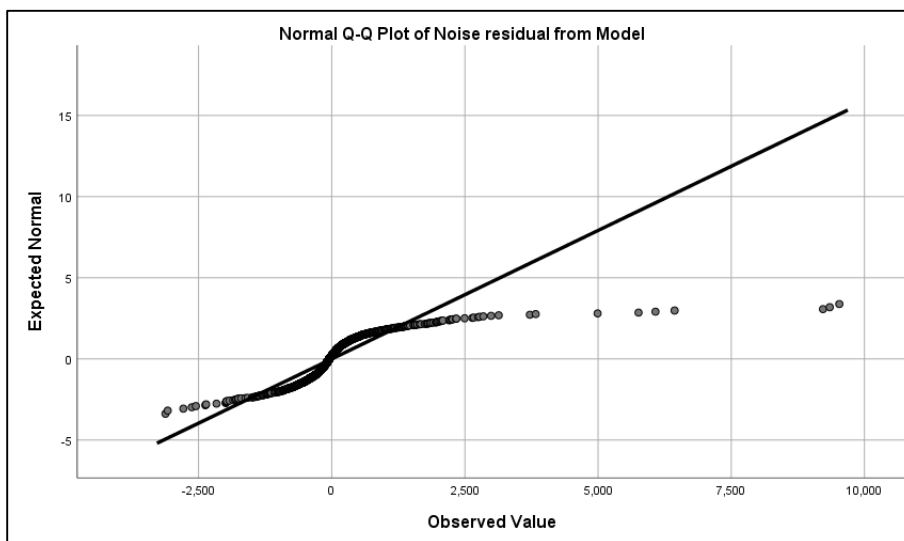
شکل آ- ۴۶ نمودار ACF و PACF برای داده‌های اصلی در محور سدا کرده - مشهد



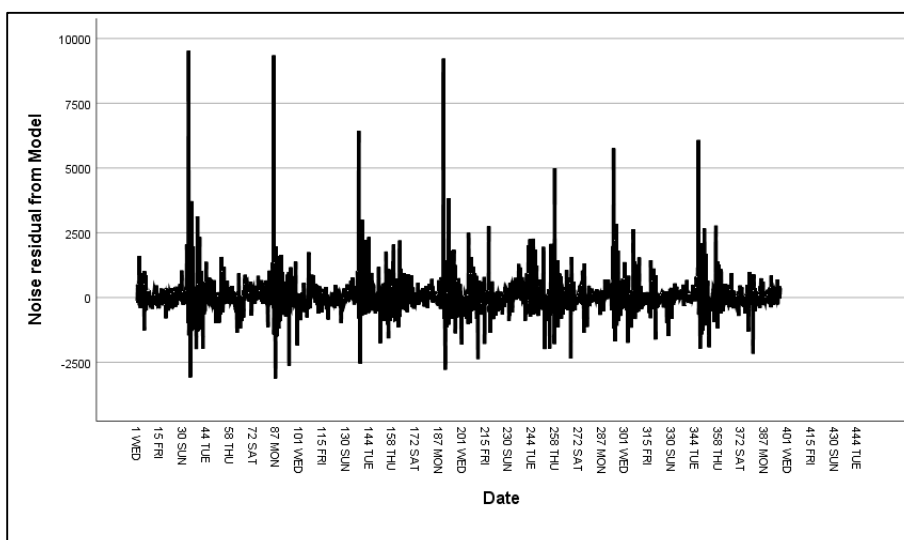
شکل آ- ۴۷ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور سدا کرده - مشهد



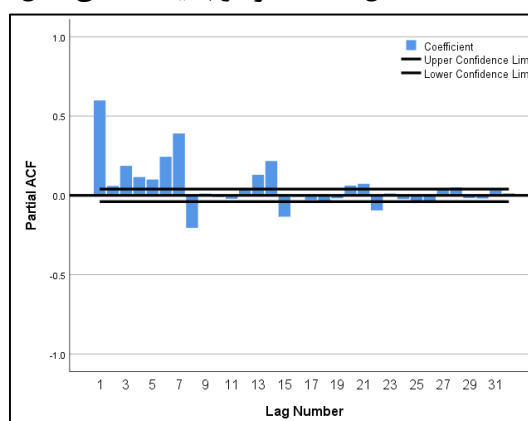
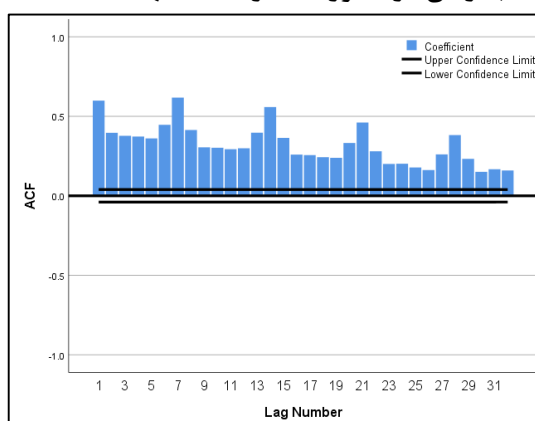
شکل آ- ۴۸ نمودارهای ACF و PACF باقیمانده‌های در محور سدا کرده - مشهد



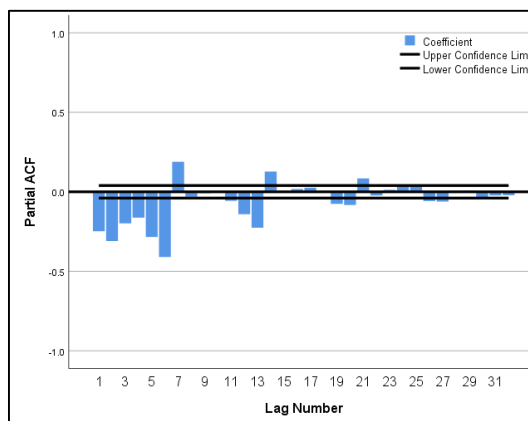
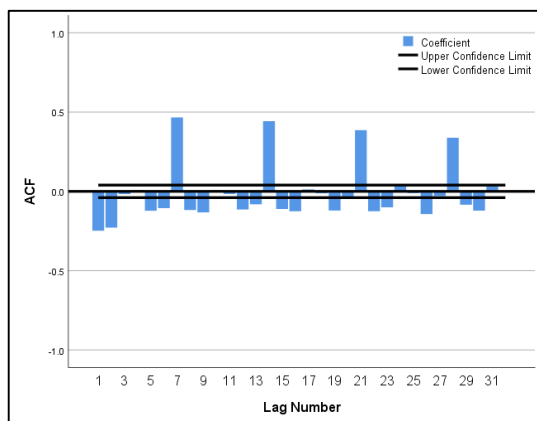
شکل آ- ۴۹ نمودار احتمال نرمال باقیمانده‌ها در محور سدکارده - مشهد



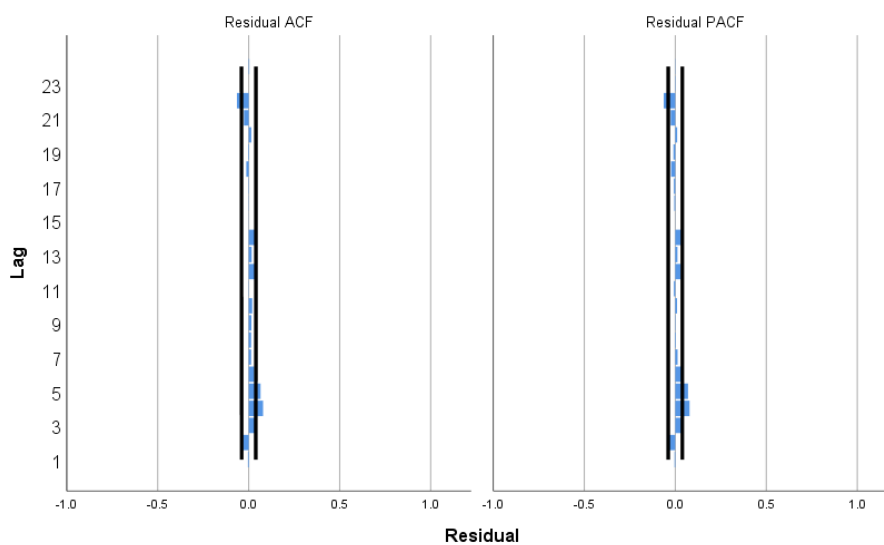
شکل آ- ۵۰ نمودار باقیمانده‌های مدل نسبت به زمان در محور سدکارده - مشهد



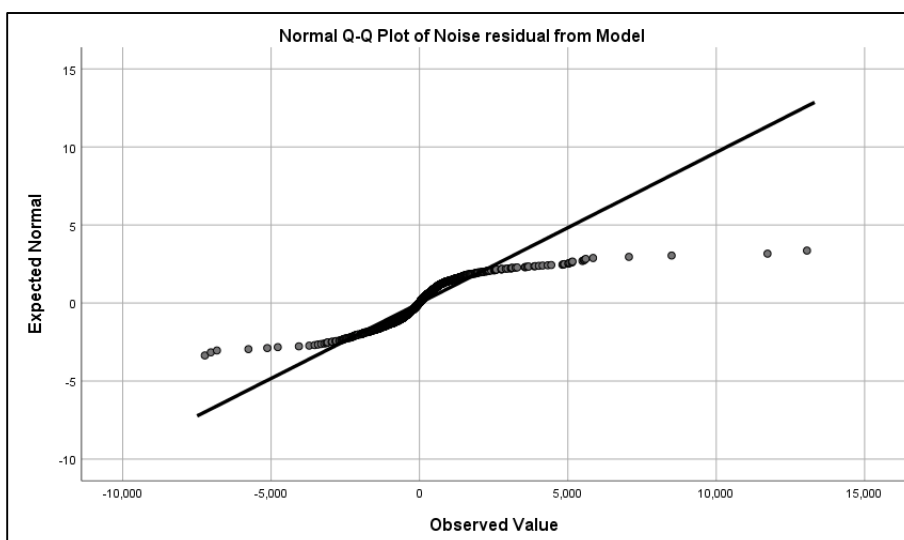
شکل آ- ۵۱ نمودار ACF و PACF برای داده‌های اصلی در محور مشهد - سدکارده



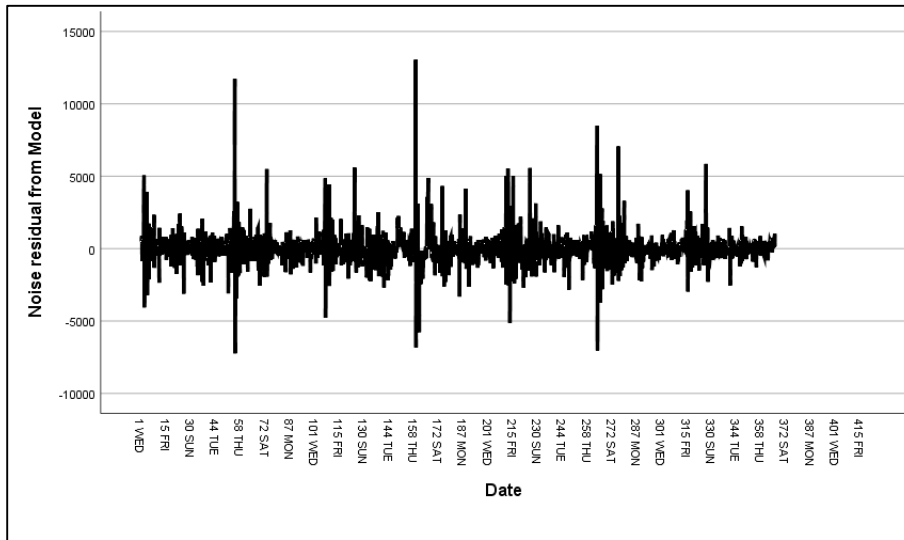
شکل آ- ۵۲ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور مشهد - سدکارده



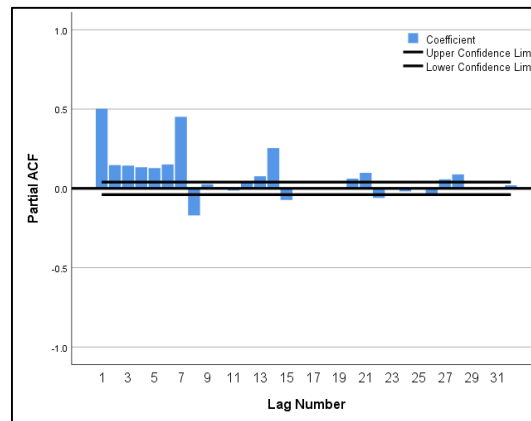
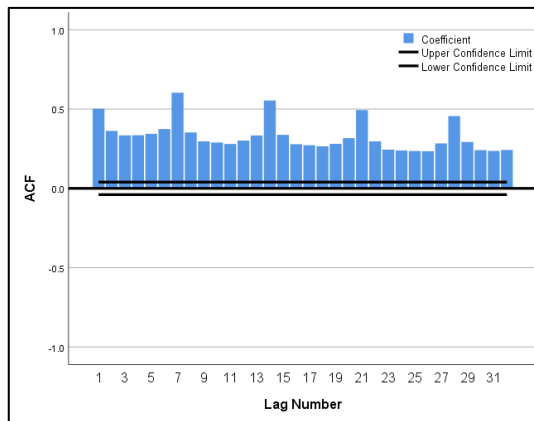
شکل آ- ۵۳ نمودارهای ACF و PACF باقیمانده‌های در محور مشهد - سدکارده



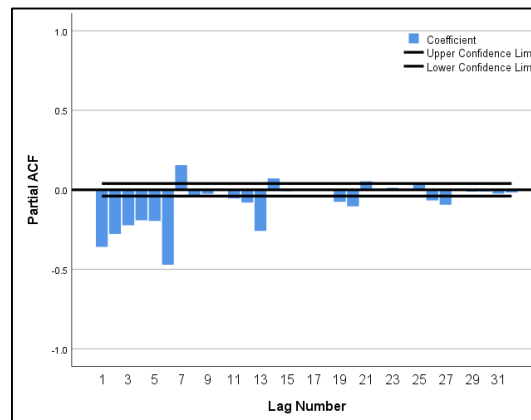
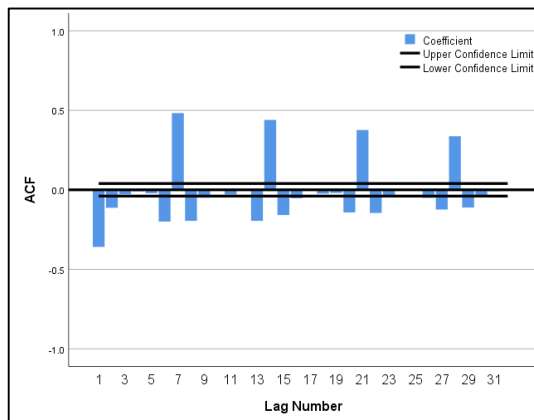
شکل آ- ۵۴ نمودار احتمال نرمال باقیمانده‌ها در محور مشهد - سدکارده



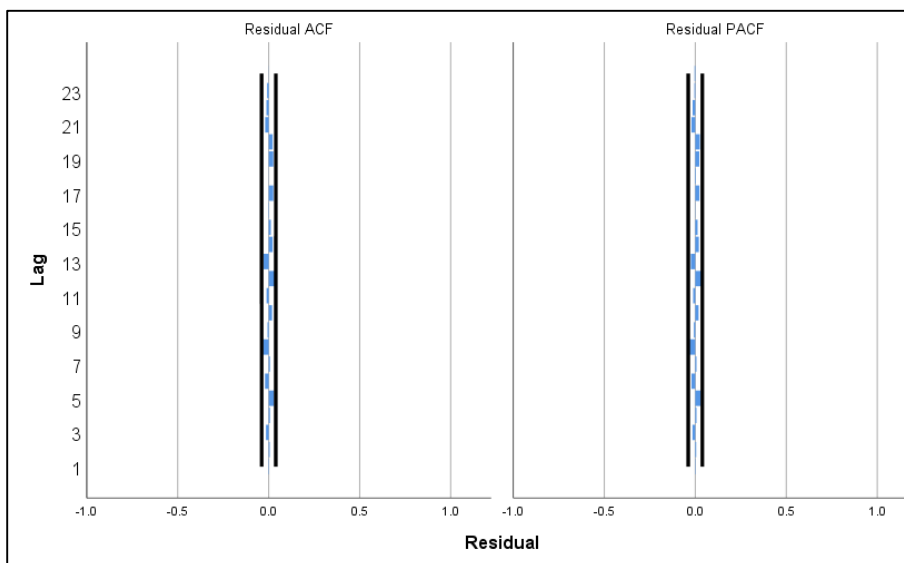
شکل آ- ۵۵ نمودار باقیمانده‌های مدل نسبت به زمان در محور مشهد - سدکارده



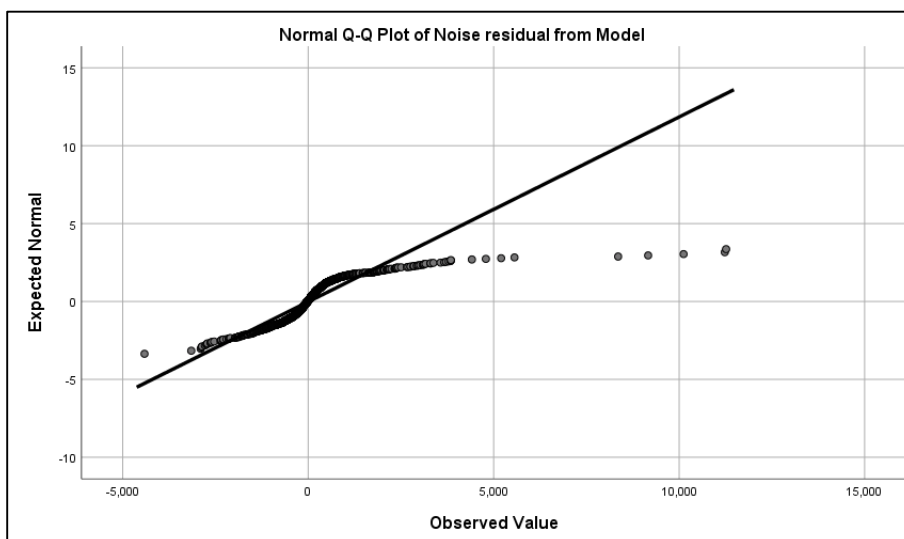
شکل آ- ۵۶ نمودار ACF و PACF برای داده‌های اصلی در محور سدکارده - مشهد



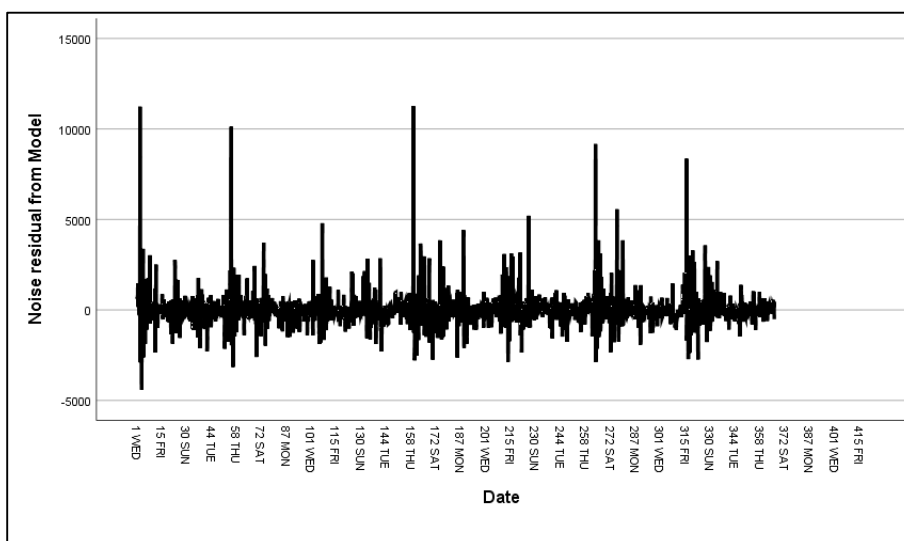
شکل آ- ۵۷ نمودار ACF و PACF برای داده‌های یکبار تفاضلی شده در محور سدکارده - مشهد



شکل آ- ۵۸ نمودارهای ACF و PACF باقیمانده‌های در محور سدا کرده - مشهد



شکل آ- ۵۹ نمودار احتمال نرمال باقیمانده‌ها در محور سدا کرده - مشهد



شکل آ- ۶۰ نمودار باقیمانده‌های مدل نسبت به زمان در محور سدا کرده - مشهد

Abstract

Traffic flow forecasting has attracted a lot in the last few decades. This is still a challenging issue for traffic and transportation researchers. Due to the random characteristics of the traffic flow, accurately predicting traffic will not be easy. Therefore, to solve this problem, many techniques have been used to model and predict traffic flow. In the present study, the parametric method, ARIMA (AutoRegressive Integrated Moving Average), non-parametric method, artificial neural network and deep learning method, LSTM (Long short-term memory) to predict the volume of road traffic It has been used for 7 round trip routes leading to Mashhad and the results of these three methods have been compared with each other. The results show that the LSTM deep learning method performs better than the other two methods. The values of RMSE, MAPE and AADT indices for these three methods are significantly different from each other. The difference between RMSE, MAPE and AADT indices of ARIMA model and LSTM deep learning model is 0.72, 24% and 23%, respectively. Also, the difference between the values of these indicators for the artificial neural network model and LSTM deep learning is 0.18, 11% and 7%, respectively, which shows the high ability of deep learning methods compared to parametric and non-parametric methods.

Keywords: Traffic volume forecasting, Time series, ARIMA, LSTM, Artificial neural network, Deep learning



**Shahrood University of
Technology**

Faculty of Civil Engineering

Master Thesis in Civil Engineering - Roads and Transportation

Traffic flow prediction using time series

By: Mojtaba Mohammadzadeh

Supervisor:
Dr. choupani

January, ۲۰۲۲